

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 29, 2005, 14:44:29 ; Search time 218 Seconds
(without alignments)
6214.849 Million cell updates/sec

Title: US-10-069-353A-7

Perfect score: 828
1 gctgtccagtggtgcgcacc.....cgcggaagccggtcgcgcga 828

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 8181359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/6A.COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/6B.COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/PTUS.COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB	ID	Description
1	828	100.0	80161	3	US-09-036-987A-1	Sequence 1, Appl1
2	828	100.0	80161	3	US-09-370-700-1	Sequence 1, Appl1
3	828	100.0	80161	4	US-09-603-207-1	Sequence 1, Appl1
4	165.4	20.0	6085	4	US-09-029-603-4	Sequence 4, Appl1
5	165.4	20.0	80161	3	US-09-036-987A-1	Sequence 1, Appl1
6	165.4	20.0	80161	3	US-09-370-700-1	Sequence 1, Appl1
7	165.4	20.0	80161	4	US-09-603-207-1	Sequence 1, Appl1
8	133.2	16.1	852	4	US-09-266-965-36	Sequence 16, Appl1
9	133.2	16.1	53500	4	US-09-266-965-76	Sequence 76, Appl1
10	122.8	14.8	2122	3	US-09-029-603-1	Sequence 1, Appl1
11	83.2	10.0	828	3	US-09-266-965-35	Sequence 35, Appl1
12	71.2	8.6	4403765	3	US-09-103-840A-2	Sequence 2, Appl1
13	71.2	8.6	4411529	3	US-09-103-840A-1	Sequence 1, Appl1
14	63.2	7.6	1369	4	US-09-439-554-29	Sequence 29, Appl1
15	62.2	7.5	10263	4	US-09-902-540-1024	Sequence 1024, Ap
16	61	7.4	1131	4	US-09-252-991A-2322	Sequence 2322, Ap
17	61	7.4	1215	4	US-09-252-991A-2513	Sequence 2513, Ap
18	61	7.4	1215	4	US-09-252-991A-2555	Sequence 2555, Ap
19	60.4	7.3	696	4	US-09-540-3933	Sequence 3933, Ap
20	57.2	6.9	991	4	US-09-439-554-27	Sequence 27, Appl1
21	57.2	6.9	1152	4	US-09-614-221A-523	Sequence 523, App
22	54.8	6.6	507	4	US-09-439-554-13	Sequence 13, Appl1
23	51.4	6.2	669	4	US-09-902-540-9153	Sequence 9153, Ap
24	51.4	6.2	17245	4	US-09-902-540-1073	Sequence 1073, Ap
25	49.4	6.0	1212	4	US-09-252-991A-9135	Sequence 9135, Ap
26	49.4	6.0	1746	4	US-09-252-991A-8948	Sequence 8948, Ap
27	49.4	6.0	2304	4	US-09-252-991A-8731	Sequence 8731, Ap

28	48.8	5.9	507	4	US-09-902-540-7484	Sequence 7484, Ap
29	48.8	5.9	6599	4	US-09-902-540-724	Sequence 724, App
30	48.4	5.8	1428	4	US-09-902-540-8152	Sequence 8152, App
31	48.4	5.8	6269	4	US-09-902-540-836	Sequence 836, App
32	47.4	5.7	2370	1	US-08-104-072B-7	Sequence 7, Appl1
33	47.4	5.7	2370	1	US-08-351-413-8	Sequence 8, Appl1
34	47.4	5.7	2370	2	US-09-025-583-8	Sequence 8, Appl1
35	47.2	5.7	1893	4	US-09-252-991A-9055	Sequence 9055, Ap
36	47.2	5.7	4403765	3	US-09-103-840A-2	Sequence 2, Appl1
37	47	5.7	4411529	3	US-09-103-840A-1	Sequence 1, Appl1
38	45	5.4	783	4	US-09-252-991A-15973	Sequence 15973, A
39	45	5.4	933	4	US-09-252-991A-16068	Sequence 16068, A
40	44.8	5.4	510	4	US-09-252-991A-8834	Sequence 8834, Ap
41	44.2	5.3	957	3	US-09-382-906A-1	Sequence 1, Appl1
42	44.2	5.3	2985	3	US-08-724-814-18	Sequence 18, Appl1
43	44.2	5.3	4632	4	US-09-902-540-8368	Sequence 8368, Ap
44	44.2	5.3	15482	4	US-09-902-540-1067	Sequence 1067, Ap
45	44	5.3	597	4	US-09-252-991A-16384	Sequence 16384, A

ALIGNMENTS

RESULT 1
US-09-036-987A-1
Sequence 1, Application US/09036987A
Patent No. 6143526
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Maddur, Krishnamurthy
APPLICANT: Merlo, Donald J.
APPLICANT: Treadway, Patti J.
APPLICANT: Turner, Jan R.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: BioSynthetic Genes For Spinosyn Insecticide
TITLE OF INVENTION: Production
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSER: Dow Agrosciences LLC Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,987A
FILING DATE: 09-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 80161 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-036-987A-1
Query Match 100.0%; Score 828; DB 3; Length 80161;
Best Local Similarity 100.0%; Pred. No. 1.7e-179;

```
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGTTCCAGAGTGGCCGCAACCATCGACAGAGTTGGGCAATGATGACCGGTGACG 60
Db 20168 GTGTTCCAGAGTGGCCGCAACCATCGACAGAGTTGGGCAATGATGACCGGTGACG 20227
QY 61 CCGTTGCTGAATCTCGGTCCGCGGGGCGCCCTGTGCGCATCCACCGGCTACTGGAGAAC 120
Db 20228 CCGTTGCTGAATCTCGGTCCGCGGGGCGCCCTGTGCGCATCCACCGGCTACTGGAGAAC 20287
QY 121 GACGGGGGGGCTTCCCTGAGAGAGAGCGCGACCGGCTACCGGCTACTTGGCCGGAACG 180
Db 20288 GACGGGGGGGCTTCCCTGAGAGAGAGCGCGACCGGCTACCGGCTACTTGGCCGGAACG 20347
QY 181 ACCGTGCTGATGAGCGGCGCTTGCATGCTGATGAGGGGTGGGTAACCGAACACGAGG 240
Db 20348 ACCGTGCTGATGAGCGGCGCTTGCATGCTGATGAGGGGTGGGTAACCGAACACGAGG 20407
QY 241 CTGCGCGTGGCGGCGCAACCGCATCCAGATCCCGGATCCGTCAGCCAGGTGCA 300
Db 20408 CTGCGCGTGGCGGCGCAACCGCATCCAGATCCCGGATCCGTCAGCCAGGTGCA 20467
QY 301 GTGGGCAATGGCCGCTGATGGCGGCAAGCGGCACTAAGCAACCGGCTGACCTTCTG 360
Db 20468 GTGGGCAATGGCCGCTGATGGCGGCAAGCGGCACTAAGCAACCGGCTGACCTTCTG 20527
QY 361 TCGCTGATGATGCTCCTGCGCTACCGGCAAAATGCTTTCGACGCGCTGGGCGATG 420
Db 20528 TCGCTGATGATGCTCCTGCGCTACCGGCAAAATGCTTTCGACGCGCTGGGCGATG 20587
QY 421 CAGTGGCTGTGAGATGTCGCAACCGGACCGGTCATCCGGAATACTTCGAGTATC 480
Db 20588 CAGTGGCTGTGAGATGTCGCAACCGGACCGGTCATCCGGAATACTTCGAGTATC 20647
QY 481 AAACCCGGGAGCATCTCGGCGTCAACCGAGTGTCAAAAGAGAGCGGCGCGGAGTG 540
Db 20648 AAACCCGGGAGCATCTCGGCGTCAACCGAGTGTCAAAAGAGAGCGGCGCGGAGTG 20707
QY 541 CCGGTGTCCGGGAGAGAGTGGCGCAACCGGCTTCGATCTGCTGAGCACTTCTG 600
Db 20708 CCGGTGTCCGGGAGAGAGTGGCGCAACCGGCTTCGATCTGCTGAGCACTTCTG 20767
QY 601 GAATGCTGCTGAGCGGCGGCTTCAGATCTTCGATTTGGAGAGAGCGTGTCTGAGAGAC 660
Db 20768 GAATGCTGCTGAGCGGCGGCTTCAGATCTTCGATTTGGAGAGAGCGTGTCTGAGAGAC 20827
QY 661 CCGTACTTCATGCGCGAGTTCGCGAAGAGCTCGGTCGCGCACGAGAGCGGAGTGC 720
Db 20828 CCGTACTTCATGCGCGAGTTCGCGAAGAGCTCGGTCGCGCACGAGAGCGGAGTGC 20887
QY 721 AGTACGGGCGGCTGTGTCGCGCTGGGCGCGCGGCTGTGCAATTATGAGAAATATGCC 780
Db 20888 AGTACGGGCGGCTGTGTCGCGCTGGGCGCGCGGCTGTGCAATTATGAGAAATATGCC 20947
QY 781 CACGACATGGGCTATGCGATTCTGACGCGCGGAGAGCGGCTGAGTGA 828
Db 20948 CACGACATGGGCTATGCGATTCTGACGCGCGGAGAGCGGCTGAGTGA 20995

RESULT 2
US-09-370-700-1
; Sequence 1, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:
; APPLICANT: Balitz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 Div1
```

```
/ CURRENT APPLICATION NUMBER: US/09/370,700
/ BEST LOCAL SIMILARITY 100.0%; DB 3; Length 80161;
/ EARLIER APPLICATION NUMBER: US 09/36987
/ EARLIER FILING DATE: 1998-03-09
/ NUMBER OF SEQ ID NOS: 39
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 1
/ LENGTH: 80161
/ TYPE: DNA
/ ORGANISM: Saccharopolyspora spinosa
US-09-370-700-1

Query Match 100.0%; Score 828; DB 3; Length 80161;
Best Local Similarity 100.0%; Pred. No. 1,7e-179;
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGTTCCAGAGTGGCCGCAACCATCGACAGAGTTGGGCAATGATGACCGGTGACG 60
Db 20168 GTGTTCCAGAGTGGCCGCAACCATCGACAGAGTTGGGCAATGATGACCGGTGACG 20227
QY 61 CCGTTGCTGAATCTCGGTCCGCGGGGCGCCCTGTGCGCATCCACCGGCTACTGGAGAAC 120
Db 20228 CCGTTGCTGAATCTCGGTCCGCGGGGCGCCCTGTGCGCATCCACCGGCTACTGGAGAAC 20287
QY 121 GACGGGGGGGCTTCCCTGAGAGAGAGCGCGACCGGCTACCGGCTACTTGGCCGGAACG 180
Db 20288 GACGGGGGGGCTTCCCTGAGAGAGAGCGCGACCGGCTACCGGCTACTTGGCCGGAACG 20347
QY 181 ACCGTGCTGATGAGCGGCGCTTGCATGCTGATGAGGGGTGGGTAACCGAACACGAGG 240
Db 20348 ACCGTGCTGATGAGCGGCGCTTGCATGCTGATGAGGGGTGGGTAACCGAACACGAGG 20407
QY 241 CTGCGCGTGGCGGCGCAACCGCATCCAGATCCCGGATCCGTCAGCCAGGTGCA 300
Db 20408 CTGCGCGTGGCGGCGCAACCGCATCCAGATCCCGGATCCGTCAGCCAGGTGCA 20467
QY 301 GTGGGCAATGGCCGCTGATGGCGGCAAGCGGCACTAAGCAACCGGCTGACCTTCTG 360
Db 20468 GTGGGCAATGGCCGCTGATGGCGGCAAGCGGCACTAAGCAACCGGCTGACCTTCTG 20527
QY 361 TCGCTGATGATGCTCCTGCGCTACCGGCAAAATGCTTTCGACGCGCTGGGCGATG 420
Db 20528 TCGCTGATGATGCTCCTGCGCTACCGGCAAAATGCTTTCGACGCGCTGGGCGATG 20587
QY 421 CAGTGGCTGTGAGATGTCGCAACCGGACCGGTCATCCGGAATACTTCGAGTATC 480
Db 20588 CAGTGGCTGTGAGATGTCGCAACCGGACCGGTCATCCGGAATACTTCGAGTATC 20647
QY 481 AAACCCGGTGCATCTTCGCGGTCAACCGAGTGTCAAAAGAGAGCGGCGCGGAGTG 540
Db 20648 AAACCCGGTGCATCTTCGCGGTCAACCGAGTGTCAAAAGAGAGCGGCGCGGAGTG 20707
QY 541 CCGGTGTCCGGGAGAGAGTGGCGCAACCGGCTTCGATCTGCTGAGCACTTCTG 600
Db 20708 CCGGTGTCCGGGAGAGAGTGGCGCAACCGGCTTCGATCTGCTGAGCACTTCTG 20767
QY 601 GAATGCTGCTGAGCGGCGGCTTCAGATCTTCGATTTGGAGAGAGCGTGTCTGAGAGAC 660
Db 20768 GAATGCTGCTGAGCGGCGGCTTCAGATCTTCGATTTGGAGAGAGCGTGTCTGAGAGAC 20827
QY 661 CCGTACTTCATGCGCGAGTTCGCGAAGAGCTCGGTCGCGCACGAGAGCGGAGTGC 720
Db 20828 CCGTACTTCATGCGCGAGTTCGCGAAGAGCTCGGTCGCGCACGAGAGCGGAGTGC 20887
QY 721 AGTACGGGCGGCTGTGTCGCGCTGGGCGCGCGGCTGTGCAATTATGAGAAATATGCC 780
Db 20888 AGTACGGGCGGCTGTGTCGCGCTGGGCGCGCGGCTGTGCAATTATGAGAAATATGCC 20947
QY 781 CACGACATGGGCTATGCGATTCTGACGCGCGGAGAGCGGCTGAGTGA 828
Db 20948 CACGACATGGGCTATGCGATTCTGACGCGCGGAGAGCGGCTGAGTGA 20995
```



```
RESULT 3
US-09-603-207-1
; Sequence 1, Application US/09603207B
; Patent No. 6521406
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 Div1
; CURRENT APPLICATION NUMBER: US/09/603,207B
; CURRENT FILING DATE: 2000-06-23
; EARLIER APPLICATION NUMBER: 09/370,700
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 80161
; TYPE: DNA
; ORGANISM: Saccharopolyspora spinosa
US-09-603-207-1

Query Match      100.0%; Score 828; DB 4; Length 80161;
Best Local Similarity 100.0%; Pred. No. 1.7e-179;
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTTGTCCAGGTGCGCCAGCAACATCGCAGAGTGGCGAGTGTATGACCTGGTCAAG 60
DB 20168 GTTGTCCAGGTGCGCCAGCAACATCGCAGAGTGGCGAGTGTATGACCTGGTCAAG 20227
QY 61 CCGTGTCTGAATCGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
DB 20228 CCGTGTCTGAATCGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20287
QY 121 GACGGGCGGGCTTCTTGAGCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
DB 20288 GACGGGCGGGCTTCTTGAGCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20347
QY 181 ACCGTGTCTGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
DB 20348 ACCGTGTCTGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20407
QY 241 CTGCGGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
DB 20408 CTGCGGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20467
QY 301 GTGGCCATCGCGCGCTGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
DB 20468 GTGGCCATCGCGCGCTGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20527
QY 361 TGGGTGATGCGCATGTCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
DB 20528 TGGGTGATGCGCATGTCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20587
QY 421 CAGTGTGTTGGAGATGTCGGAACCGGACCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
DB 20588 CAGTGTGTTGGAGATGTCGGAACCGGACCGTGGCGGCGGCGGCGGCGGCGGCGGCGG 20647
QY 481 AAAACCCGGGTGCGATCTCTGCGGCGTCAACGAGGTGCTCAACGAGAGCGGCGGCGGCGG 540
DB 20648 AAAACCCGGGTGCGATCTCTGCGGCGTCAACGAGGTGCTCAACGAGAGCGGCGGCGGCGG 20707
QY 541 CCGGTGTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
DB 20708 CCGGTGTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20767
QY 601 GAATGCGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
DB 20768 GAATGCGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20827
```

```
QY 661 CGGTACTTCATGCGCGAGTTCGCGGAGAGCTGCTGCGGCAACGAGCGGATCCGCGAC 720
DB 20828 CGGTACTTCATGCGCGAGTTCGCGGAGAGCTGCTGCGGCAACGAGCGGATCCGCGAC 20887
QY 721 AGTACCGGCGCGGCTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
DB 20888 AGTACCGGCGCGGCTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20947
QY 781 CACGACATGCGGCTATGCGATTTGACCGGCGGCGGAGCCGCGTGGCTGA 828
DB 20948 CACGACATGCGGCTATGCGATTTGACCGGCGGCGGAGCCGCGTGGCTGA 20995

RESULT 4
US-09-029-603-4
; Sequence 4, Application US/09029603
; Patent No. 6210935
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Engel, Natalie
; APPLICANT: Bielenhader, Jurg
; APPLICANT: Toupet, Christine
; APPLICANT: Pospishech, Andreas
; TITLE OF INVENTION: Staurosporin Biosynthesis Gene Clusters
; FILE REFERENCE: 4-20555/A/PCT
; CURRENT APPLICATION NUMBER: US/09/029,603
; CURRENT FILING DATE: 1998-03-20
; EARLIER APPLICATION NUMBER: PCT/EP96/03643
; EARLIER FILING DATE: 1996-08-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 6085
; TYPE: DNA
; ORGANISM: Streptomyces longisporoflavus
; FEATURE:
; NAME/KEY: misc RNA
; LOCATION: (378)..(1665)
; OTHER INFORMATION: ORF
; FEATURE:
; NAME/KEY: misc RNA
; LOCATION: (1747)..(2553)
; OTHER INFORMATION: ORF
; FEATURE:
; NAME/KEY: misc RNA
; LOCATION: (2593)..(4011)
; OTHER INFORMATION: ORF
; FEATURE:
; NAME/KEY: misc RNA
; LOCATION: (4013)..(4999)
; OTHER INFORMATION: ORF
; NAME/KEY:
; NAME/KEY: misc RNA
; LOCATION: (5071)..(6085)
; OTHER INFORMATION: ORF
US-09-029-603-4

Query Match      20.0%; Score 165.4; DB 3; Length 6085;
Best Local Similarity 60.6%; Pred. No. 1.6e-28;
Matches 271; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 90 CTGCGCATTCACACAGGCGTATCTGGAGAGAACGAGCGGCGGCTTCTGCGGACGAGCGCG 149
DB 1833 CGGCAACATTCACAGGCGTATCTGGAGAGAACGAGCGGCGGCTTCTGCGGACGAGCGCG 1892
QY 150 CGACCGGCTACCGGACCTTGTGCGCGGAGCGGACCGGTGCTGATGCGGCGGCGGCTTGA 209
DB 1893 CGACCGGCTACCGGACCTTGTGCGCGGAGCGGACCGGTGCTGATGCGGCGGCGGCTTGA 1952
QY 210 CGATGTGGGCGGTGATCCGAGCAACGAGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGG 269
DB 1953 GGAAGTGGGCGGTGATCCGAGCAACGAGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGG 2012
```


SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 1
 LENGTH: 80161
 TYPE: DNA
 ORGANISM: Saccharopolyspora spinosa
 US-09-370-700-1

Query Match 20.0%; Score 165.4; DB 3; Length 80161;
 Best Local Similarity 52.8%; Pred. No. 2.3e-28;
 Matches 386; Conservative 0; Mismatches 336; Indels 9; Gaps 1;

```

QY 15 CGCACCAACATCGCAGAGTTGGGAGATGTATGACCTGTGCAACCGCTGTGTAATC 74
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13575 CACACAGGATTTCCGACCGCGGATCAGGTGAGTCACTTGTGAGCGGACCG 13516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 75 GGTCCGGGGGCGCCCTGCGCATCCACACCGCTTACGTGGAGAACGACGGGCGGCTTC 134
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13515 GGTCCCTCCGACACAGGTTACTGGGCGGGGTTACGGGAGGATGCGGTGCCACACC 13456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 135 CTGGCAGCAGCGCGCGGACCGGCTCACCGACTTGTGCGGAAACGACCGTGTGATGG 194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13455 GTGTGCGATGTCTGCGGACCACTGACCGACTGTTCATGACAAAGCGCGCTCGTCC 13396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 195 CGGCGTTGACCTGCTGATGTGGGGTGGCGGTACCGGACACAGCGCTGCGCTCGCGG 254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13395 CGGAGCGCACCTGTTGACCTGGGCTGGCGGATGGGCGACCCGTAGTCCGTGGCGATG 13336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 255 CGACACCGGATCCAGATCACCGGACATCACTGACCGAGTCCAGAGTCCATGCGCGG 314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13335 CGGACAGGGGTTTCCAGATCACCGGAATCACCGTGAAGCCGACGACTCTGCGCGCCAC 13276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 315 TGATTGCGCAGCGGAAACCGGAACTTAAGCCACCGGGTGGACTTTCGTGCTGATGCCAT 374
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13275 CAGGCTGCGCAACGAGACCGGACTGGCCGACGTGATGATCTAGTTCAGCGCGC 13216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 375 GTCCCTGCGTACCGGAGCAATGCTTTCAGACCGCGCTGGGCAATGCTGTTTGA 434
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13215 CAGCTGCGCTTACCGGACCGGCTTCTTTCAGGCGGATGGGCAATGATGCTGCTGCA 13156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 435 GATGTCGGAACCGGACCGGTGCCATCCGGGAATCTTTCAGATCTCAAAACCGGTGGCAT 494
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13155 GATGTCGGAACCGGACCGGTGCCATCCGGGAATCTTTCAGATCTCAAAACCGGTGGCAT 13096
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 495 CTTGCGCGTCAACGAGGTCTCAACGAGAACGGGGGCGGGATGCGGTGTCCGGGCA 554
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13095 GTTCTGCTCGAGACATCATCTCGGGTTCGACTCCCGAAGATGACCGCGGCTTGG 13036
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 555 CAGGTGGCGGACCGGCTTTCGATCTGCGCTGGAGCACTTCTGGAATGCTGGGCTG 614
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13035 GACGGGACGACCGCCCATACCTT-----GAAACGCTTCAAGCGCTGTGACGCA 12985
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 615 AGCGGGTTTGAATCTCGATTGGAGAGACGTGTGTCGAGAACCGGATCTTCAATGCC 674
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 12984 AGCGGGTTTGAATCTCGATTGGAGAGACGTGTGTCGAGAACCGGATCTTCAATGCC 12925
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 675 GCAGTTGCGGAAAGCTCGCTGGGACACGACCGGATGCGGACGGTACGGGCGGCG 734
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 12924 CTGTGATGTCGAGAGATGCTCGGAAACTGATGAGCTGCGGCGCTGAGGCTGCGG 12865
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 735 TGTGCGCGGCT 745
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 12864 TGTGCGCACCT 12854
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 7
 US-09-603-207-1/C
 Sequence 1, Application US/09603207B
 Patent No. 6521406
 GENERAL INFORMATION:
 APPLICANT: Baltz, Richard H
 APPLICANT: Broughton, Mary C
 APPLICANT: Crawford, Kathryn P
 APPLICANT: Madduri, Krishnamurthy

APPLICANT: Treadway, Patti J
 APPLICANT: Turner, Jan R
 APPLICANT: Waldron, Clive
 TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
 FILE REFERENCE: 50489 DIV1
 CURRENT APPLICATION NUMBER: US/09/603,207B
 EARLIER FILING DATE: 2000-06-23
 EARLIER APPLICATION NUMBER: 09/370,700
 NUMBER OF SEQ ID NOS: 39
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 1
 LENGTH: 80161
 TYPE: DNA
 ORGANISM: Saccharopolyspora spinosa
 US-09-603-207-1

Query Match 20.0%; Score 165.4; DB 4; Length 80161;
 Best Local Similarity 52.8%; Pred. No. 2.3e-28;
 Matches 386; Conservative 0; Mismatches 336; Indels 9; Gaps 1;

```

QY 15 CGCACCAACATCGCAGAGTTGGGAGATGTATGACCTGTGCAACCGCTGTGTAATC 74
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13575 CACACAGGATTTCCGACCGCGGATCAGGTGAGTCACTTGTGAGCGGACCG 13516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 75 GGTCCGGGGGCGCCCTGCGCATCCACACCGCTTACGTGGAGAACGACGGGCGGCTTC 134
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13515 GGTCCCTCCGACACAGGTTACTGGGCGGGGTTACGGGAGGATGCGGTGCCACACC 13456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 135 CTGGCAGCAGCGCGCGGACCGGCTCACCGACTTGTGCGGAAACGACCGTGTGATGG 194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13455 GTGTGCGATGTCTGCGGACCACTGACCGACTGTTCATGACAAAGCGCGCTCGTCC 13396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 195 CGGCGTTGACCTGCTGATGTGGGGTGGCGGTACCGGACACAGCGCTGCGCTCGCGG 254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13395 CGGAGCGCACCTGTTGACCTGGGCTGGCGGATGGGCGACCCGTAGTCCGTGGCGATG 13336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 255 CGACACCGGATCCAGATCACCGGACATCACTGACCGAGTCCAGAGTCCATGCGCGG 314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13335 CGGACAGGGGTTTCCAGATCACCGGAATCACCGTGAAGCCGACGACTCTGCGCGCCAC 13276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 315 TGATTGCGCAGCGGAAACCGGAACTTAAGCCACCGGGTGGACTTTCGTGCTGATGCCAT 374
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13275 CAGGCTGCGCAACGAGACCGGACTGGCCGACGTGATGATCTAGTTCAGCGCGC 13216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 375 GTCCCTGCGTACCGGAGCAATGCTTTCAGACCGCGCTGGGCAATGCTGTTTGA 434
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13215 CAGCTGCGCTTACCGGACCGGCTTCTTTCAGGCGGATGGGCAATGATGCTGCTGCA 13156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 435 GATGTCGGAACCGGACCGGTGCCATCCGGGAATCTTTCAGATCTCAAAACCGGTGGCAT 494
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13155 GATGTCGGAACCGGACCGGTGCCATCCGGGAATCTTTCAGATCTCAAAACCGGTGGCAT 13096
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 495 CTTGCGCGTCAACGAGGTCTCAACGAGAACGGGGGCGGGATGCGGTGTCCGGGCA 554
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13095 GTTCTGCTCGAGACATCATCTCGGGTTCGACTCCCGAAGATGACCGCGGCTTGG 13036
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 555 CAGGTGGCGGACCGGCTTTCGATCTGCGCTGGAGCACTTCTGGAATGCTGGGCTG 614
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13035 GACGGGACGACCGCCCATACCTT-----GAAACGCTTCAAGCGCTGTGACGCA 12985
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 615 AGCGGGTTTGAATCTCGATTGGAGAGACGTGTGTCGAGAACCGGATCTTCAATGCC 674
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 12984 AGCGGGTTTGAATCTCGATTGGAGAGACGTGTGTCGAGAACCGGATCTTCAATGCC 12925
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 675 GCAGTTGCGGAAAGCTCGCTGGGACACGACCGGATGCGGACGGTACGGGCGGCG 734
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 12924 CTGTGATGTCGAGAGATGCTCGGAAACTGATGAGCTGCGGCGCTGAGGCTGCGG 12865
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 735 TGTGCGCGGCT 745
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 12864 TGTGCGCACCT 12854
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 8

US-09-266-965-36
Sequence 36, Application US/09266965

Patent No. 6495348

GENERAL INFORMATION:

APPLICANT: Sherman, D

APPLICANT: Mao, Y

APPLICANT: Varoglu, M

APPLICANT: He, M

APPLICANT: Sheldon, P

TITLE OF INVENTION: Mitomycin biosynthetic gene cluster

FILE REFERENCE: 600,456U1

CURRENT APPLICATION NUMBER: US/09/266,965

EARLIER FILING DATE: 1999-03-12

EARLIER APPLICATION NUMBER: US 08/624,447

EARLIER FILING DATE: 1996-08-19

EARLIER APPLICATION NUMBER: PCT/US94/11279

EARLIER FILING DATE: 1994-10-06

EARLIER APPLICATION NUMBER: US 08/133,963

EARLIER FILING DATE: 1993-10-07

NUMBER OF SEQ ID NOS: 145

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 36

LENGTH: 852

TYPE: DNA

ORGANISM: Streptomyces lavendulae

US-09-266-965-36

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

Score

DB 4

Length

Pred. No.

Indels

Gaps

Score

DB 4

Length

Pred. No.

Indels

Gaps

Score

DB 4

Length

Pred. No.

Indels

Gaps

Score

DB 4

Length

Pred. No.

Indels

Gaps

Score

DB 4

Length

Pred. No.

Indels

Gaps

Score

DB 4

Length

Pred. No.

Indels

Gaps

Score

RESULT 9

US-09-266-965-76/c
Sequence 76, Application US/09266965

Patent No. 6495348

GENERAL INFORMATION:

APPLICANT: Sherman, D

APPLICANT: Mao, Y

APPLICANT: Varoglu, M

APPLICANT: He, M

APPLICANT: Sheldon, P

TITLE OF INVENTION: Mitomycin biosynthetic gene cluster

FILE REFERENCE: 600,456U1

CURRENT APPLICATION NUMBER: US/09/266,965

EARLIER FILING DATE: 1999-03-12

EARLIER APPLICATION NUMBER: US 08/624,447

EARLIER FILING DATE: 1996-08-19

EARLIER APPLICATION NUMBER: PCT/US94/11279

EARLIER FILING DATE: 1994-10-06

EARLIER APPLICATION NUMBER: US 08/133,963

EARLIER FILING DATE: 1993-10-07

NUMBER OF SEQ ID NOS: 145

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 76

LENGTH: 53500

TYPE: DNA

ORGANISM: Streptomyces lavendulae

US-09-266-965-76

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

Score

DB 4

Length

Pred. No.

Indels

Gaps

Score

DB 4

Length

Pred. No.

Indels

Gaps

Score

DB 4

Length

Pred. No.

Indels

Gaps

Score

DB 4

Length

Pred. No.

Indels

Gaps

Score

DB 4

Length

Pred. No.

Indels

Gaps

Score

DB 4

Length

Pred. No.

Indels

Gaps

Score

QY 486 CGGTGGCATCTTCGGCGTCAACCGAGTCT 515
|||
Db 477 GGGCGCGCGGTGACGAGTACGAGCTTCT 506
|||

RESULT 12

US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 8.6%; Score 71.2; DB 3; Length 4403765;
Best Local Similarity 50.6%; Pred. No. 1.2e-06;
Matches 172; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

QY 210 CGATGTGGGGTGGTACCGGACCAACGCGCTGGCGCCCGGACCAAGCATCA 269
|||
Db 4172227 CGACATCGCGCGCGCGCGCGCGCTCGGACGCGCTGGCGCACTAGGTGTGAC 4172286
|||
QY 270 GATCAACCGGATCATCGGTACGAGGATGAGTGCATCGCGCTGATTTGGACAGCA 329
|||
Db 4172287 TGTCAACGGGGTGCATCATGAGGATCAACTGAAGCGCGCCCAAGCCGCTGGCGCA 4172346
|||
QY 330 ACGCGGACTAAGCCACCGGGTGAAGTCTTCGTCGTCGATGTCATGCTCCGCTACCC 389
|||
Db 4172347 AACCGGCTGACCGGAGCGGGTGCCTTCCACTGGGCGACGCGGATCATGCTCCGTTGCC 4172406
|||
QY 390 GGAACAATGCTTTCGACGCGCGCTGGGCAATGATGCTGTGGAGATGTCGAAACCGGA 449
|||
Db 4172407 CGACGACACATTCGACGCGCTGGTGTGAGTGCAGCGCTTCGACATTCGCGGACAGAA 4172466
|||
QY 450 CCGTCCATTCGCGGGAATCTTTCGAGTACTCAACCGGATGATCTCGGCGTACCGCA 509
|||
Db 4172467 CGCGCGCGCGCGGAGCTTGTCTCGATTCGCTGTCTGTGCTGTGGCGCGGATACCGCA 4172526
|||
QY 510 GGTGTCAACGAGAGGCGGCGGCGGAGTGGCGGTGCC 549
|||
Db 4172527 TGTCACTGTGCGGAGCGGCGGCTGCGCGGAGCTGACC 4172566
|||

RESULT 13

US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A

; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 8.6%; Score 71.2; DB 3; Length 4411529;
Best Local Similarity 50.6%; Pred. No. 1.2e-06;
Matches 172; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

QY 210 CGATGTGGGGTGGGTAACCGGACCAACGCGCTGGCGCCCGGACCAAGCATCA 269
|||
Db 4179979 CGACATCGCGCGCGCGCGCGCGCTCGGACGCGCTGTGGCGCACTACGCTGTGCGC 4180038
|||
QY 270 GATCAACCGGATCATCGGTACGAGTGCAGAGTGCATGCGCGCTGATTTGGACAGCGCA 329
|||
Db 4180039 TGTCAACGGGGTGCATCATGAGTGCAGATCACTGAAGCGCGCCCAAGCGCGCTGCGCA 4180098
|||
QY 330 ACGGGAATAAGCCACCGGGTGAAGTCTTCGTCGTCGATGATGTCATGTCCTGCGTACCC 389
|||
Db 4180099 AACCGGCTGACCGAGCGGGTGCCTTTCACCTGGGCGACGCGGATCATGCTCCGTTGCC 4180158
|||
QY 390 GGAACAATGCTTTCGACGCGCGCTGGGCAATGATGCTGTGGAGATGTCGAAACCGGA 449
|||
Db 4180159 CGACGACACATTCGACGCGCTGGTGTGAGTGCAGTGGCGCTTTCGACATTCGCGGACAGAA 4180218
|||
QY 450 CCGTCCATTCGCGGGAATCTTTCGAGTACTCAACCGGATGATCTCGGCGTACCGCA 509
|||
Db 4180219 CGCGCGCGCGGACGAGTTCGCTCGGATTCGCTGTGTCGCTGGCGCGGATCACCGA 4180278
|||
QY 510 GGTGTCAACGAGAGCGGCGGCGGAGTGGCGGTGCC 549
|||
Db 4180279 TGTCACTGTGCGGAGCGGCGGCTGCGCGGAGCTGACC 4180318
|||

RESULT 14

US-09-439-554-29
; Sequence 29, Application US/09439554
; Patent No. 6479733
; GENERAL INFORMATION:
; APPLICANT: Rafalski, Jan Antoni
; APPLICANT: Odell, Joan T.
; APPLICANT: Sakai, Hajime
; APPLICANT: Thorpe, Catherine J.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Farnodu, Omolayo O.
; TITLE OF INVENTION: STEROL METABOLISM ENZYMES
; FILE REFERENCE: BB1114 US NA
; CURRENT APPLICATION NUMBER: US/09/439,554
; CURRENT FILING DATE: 1999-11-12
; EARLIER APPLICATION NUMBER: 60/108,351
; EARLIER FILING DATE: 1998-No. 6479733ember-13
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 1369
; TYPE: DNA
; ORGANISM: Glycine max
US-09-439-554-29

Query Match 7.6%; Score 63.2; DB 4; Length 1369;
Best Local Similarity 47.3%; Pred. No. 2.5e-05;
Matches 191; Conservative 0; Mismatches 213; Indels 0; Gaps 0;

QY 133 TCTGGCAGAGGCGCGCGGACCGGCTCACCACTTTCGCGGAGGAGCGATGCTGAT 192
|||
Db 386 TCCACCGCGAGGACGCGGCTTTCACGAGAGATGCGCTTCATTCATCGAGGCCAAG 445
|||
QY 193 GCGCGGCTTGCATGCTGATGTGGGGTGGGTAACCGGACCAACGCGCTGCGGCTGCGC 252
|||

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 29, 2005, 13:43:45 ; Search time 3886 Seconds
(without alignments)
10324.477 Million cell updates/sec

Title: US-10-069-353A-7

Perfect score: 828
Sequence: 1 ggtgtccagcggcggtcgccacc.....cgcggaagccggtcggtcga 828

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmb1:*
1: gb_ba:*
2: gb_bt:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sta:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	828	100.0	828	6	AX089422 Sequence
2	828	100.0	45624	6	AX089419 Sequence
3	828	100.0	50000	6	AX089416 Sequence
4	828	100.0	80161	1	AY007564 Saccharop
5	828	100.0	80161	6	AR165018 Sequence
6	828	100.0	80161	6	BD137649 Biosynthe
7	828	100.0	80161	6	AR281866 Sequence
8	749.6	90.5	36538	6	AX600587 Sequence
9	170	20.5	846	6	AX598613 Sequence
10	170	20.5	52101	6	AX598593 Sequence
11	165.4	20.0	900	6	AX089434 Sequence
12	165.4	20.0	6085	6	AX0304 Sequence
13	165.4	20.0	6085	6	AR144763 Sequence
14	165.4	20.0	45624	6	AX089419 Sequence
15	165.4	20.0	50000	6	AX089416 Sequence
16	165.4	20.0	80161	1	AY007564 Saccharop
17	165.4	20.0	80161	6	AR165018 Sequence
18	165.4	20.0	80161	6	BD137649 Biosynthe
19	165.4	20.0	80161	6	AR281866 Sequence

20	159	19.2	36538	6	AX600587 Sequence
21	158.8	19.2	22999	1	AB088119 Streptomycin
22	142.2	17.2	210614	1	AB088224 Streptomycin
23	141.8	17.1	17512	1	AB071405 Lechevali
24	141.8	17.1	25681	1	SAE414559 Saccharo
25	141.8	17.1	26144	1	AB090952 Lechevali
26	141.8	17.1	28654	1	AF534707 Lechevali
27	140.8	17.0	846	1	AY396042 Streptomycin
28	140.6	17.0	109528	1	AF040570 Amycolato
29	137.6	16.6	5998	1	AB032524 Streptomycin
30	137.6	16.6	302875	1	AP050524 Streptomycin
31	133.2	16.1	852	6	BD269407 Streptomycin
32	133.2	16.1	852	6	AR266893 Streptomycin
33	133.2	16.1	53500	6	BD269447 Streptomycin
34	133.2	16.1	53500	6	AR266893 Streptomycin
35	133.2	16.1	63734	1	AF127374 Streptomycin
36	130	15.7	1410	6	E07846 DNA sequence
37	122.8	14.8	2122	6	A60301 Sequence 1
38	122.8	14.8	2122	6	AR144762 Sequence
39	120.2	14.5	30000	6	AX50261 Sequence
40	120.2	14.5	103450	1	AF440781 Streptomycin
41	112.2	13.6	12381	6	BD217469 Polyketid
42	112.2	13.6	12381	6	AX006889 Sequence
43	97.4	11.8	110000	8	CR382132_12 Confination (13 o
44	95	11.5	12637	1	AF323753 Streptomycin
45	93.8	11.3	1585	8	AK060887 Oryza sat

ALIGNMENTS

RESULT 1	AX089422	828 bp	DNA	linear	PAT 21-MAR-2001
LOCUS	Sequence	7 from Patent WO0116303.			
DEFINITION	AX089422				
ACCESSION	AX089422.1	GI:1343683			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM	Saccharopolyspora spinosa				
	Saccharopolyspora spinosa				
	Bacteria; Actinobacteria; Actinomycetales;				
	Pseudonocardineae; Pseudonocardaceae; Saccharopolyspora.				
REFERENCE	1				
AUTHORS	Eberz, G., Moehrl, V., Froede, R., Velten, R. and Salas, J.A.				
TITLE	Nucleic acids which code for the enzyme activities of the spinosyn				
JOURNAL	Biosynthesis				
	Patent: WO 0116303-A 7 08-MAR-2001;				
FEATURES					
source	BAYER AG (DE)				
	Location/Qualifiers				
	1..828				
	/organism="Saccharopolyspora spinosa"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:60894"				
	1..828				
	/note="unassigned protein product; ORF1, O-Methyltransferase"				
	/codon_start=1				
	/transl_table=11				
	/protein_id="CAC34874.1"				
	/db_xref="GI:1343684"				
	/translation="MLPGAGPTSQVGVQVYDVTPLNSVAGPCAIIHGVENDGRA				
	SWOQADRLTDLVAERTVLDGVRLLDVCCTGCGPALRARNALQITITVSQOVA				
	IAADCARERGLSHRVDFSCVDAMSLPYPNARFDAMAMSLIEMSEPRDAIEIRLV				
	KRGIGLCTEVYKRRAGGMPYSGRWPTGLAICLAEOLESLRAAGFEILLMEVSS				
	KRTVMPQFAEBLAAHQGIAIDRYGPAVAGMAVCDIEKXAHMDGKATITLRKPVG"				
ORIGIN					
Query Match	100.0%;	Score 828;	DB 6;	Length 828;	
Best Local Similarity	100.0%;	Pred. No. 2.4e-116;			
Matches	828;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GTTGTCAGAGTGGCGCACCAATCGACAGAGTTGGCAGATGTATGACCTGTGACG	60		
Db	1	GTTGTCAGAGTGGCGCACCAATCGACAGAGTTGGCAGATGTATGACCTGTGACG	60		

QY 61 CCGTTGCTGAACCTCGGTCCGGGCGGCCCCCTGCGCCATCCACACGGCTACTGGAGAAC 120
DB 61 CCGTTGCTGAACCTCGGTCCGGGCGGCCCCCTGCGCCATCCACACGGCTACTGGAGAAC 120
QY 121 GACGGGCGGGCTTCTCTGAGACAGAGCGCGCCGACCGGCTCACCGACCTTGTGCGCGAACG 180
DB 121 GACGGGCGGGCTTCTCTGAGACAGAGCGCGCGCCGACCGGCTCACCGACCTTGTGCGCGAACG 180
QY 181 ACCGTCCTGATGAGCGGCGCTTTCGACTGCTCGATGAGGAGGTGCGGTACCGGACAAACAGCG 240
DB 181 ACCGTCCTGATGAGCGGCGCTTTCGACTGCTCGATGAGGAGGTGCGGTACCGGACAAACAGCG 240
QY 241 CTGCGCGTCGCGCGACAAACCGGATCCAGATCACCGGATCACCGTACCGAGGTCGAA 300
DB 241 CTGCGCGTCGCGCGACAAACCGGATCCAGATCACCGGATCACCGTACCGAGGTCGAA 300
QY 301 GTGGCCATCGCCGCTGATGTCGACGCGAACGCGGACTTAAGCCACCGGGTGAATTTCTCG 360
DB 301 GTGGCCATCGCCGCTGATGTCGACGCGAACGCGGACTTAAGCCACCGGGTGAATTTCTCG 360
QY 361 TGGCGTCGATGCGATGTCCTGCGCTGACCGGACAAATGCTTTGACGCGCGCTGGGCGCATG 420
DB 361 TGGCGTCGATGCGATGTCCTGCGCTGACCGGACAAATGCTTTGACGCGCGCTGGGCGCATG 420
QY 421 CAGTCGCTGTTGAGAGATGTCGACACCGGACCGTGCATCCGGGAAATCTTTCGAGTACTC 480
DB 421 CAGTCGCTGTTGAGAGATGTCGACACCGGACCGTGCATCCGGGAAATCTTTCGAGTACTC 480
QY 481 AAACCCGGTGCATCTCTCGGCGCTCACCGAGGTGTCGAAAGAGAGGCGGCGGCGGATG 540
DB 481 AAACCCGGTGCATCTCTCGGCGCTCACCGAGGTGTCGAAAGAGAGGCGGCGGCGGATG 540
QY 541 CCGGTGTCGCGGAGACAGGTGCGCGACCGGCGCTTCCGATGCTGCGTGAAGAACTTCTCG 600
DB 541 CCGGTGTCGCGGAGACAGGTGCGCGACCGGCGCTTCCGATGCTGCGTGAAGAACTTCTCG 600
QY 601 GAATGCTGCTGACAGCGGCGGTTTCGAGATCTTCGATTTGGAGAGACGTGCTGCGAGAAC 660
DB 601 GAATGCTGCTGACAGCGGCGGTTTCGAGATCTTCGATTTGGAGAGACGTGCTGCGAGAAC 660
QY 661 CGGTACTTCATGCGCGAGATTCCGCGAAGAGCTGCTGCGCACGACGCGGATCGCGAGC 720
DB 661 CGGTACTTCATGCGCGAGATTCCGCGAAGAGCTGCTGCGCACGACGCGGATCGCGAGC 720
QY 721 AGGTACGGGCGCGCTGTCCCGCGCTGAGGCGCGCGCTGTGCGATTTAAGAAATATGCGC 780
DB 721 AGGTACGGGCGCGCTGTCCCGCGCTGAGGCGCGCGCTGTGCGATTTAAGAAATATGCGC 780
QY 781 CACGACATGGGCTATGCGATTCTGACGCGCGCGGAAACCGGTGCGCTGA 828
DB 781 CACGACATGGGCTATGCGATTCTGACGCGCGCGGAAACCGGTGCGCTGA 828

RESULT 2
AX089419/c 45624 bp DNA linear PAT 21-MAR-2001
LOCUS AX089419
DEFINITION Sequence 4 from Patent WO0116303.
ACCESSION AX089419
VERSION AX089419.1 GI:13443680
KEYWORDS
SOURCE .
ORGANISM Saccharopolyspora spinosa
Saccharopolyspora spinosa
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Pseudonocardiales; Pseudonocardaceae; Saccharopolyspora.
REFERENCE
1 Eberz, G., Moehrle, V., Froede, R., Veltan, R. and Salas, J.A.
Nucleic acids which code for the enzyme activities of the spinosyn
biosynthesis
Patent: WO 0116303-A 4 08-MAR-2001;
JOURNAL BAYER AG (DE)
FEATURES
source location/Qualifiers
1. .45624

ORIGIN
Query Match 100.0%; Score 828; DB 6; Length 45624;
Best Local Similarity 100.0%; Pred. No. 9, se-117;
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/organism="Saccharopolyspora spinosa"
/mol_type="unassigned DNA"
/db_xref="taxon:50894"

QY 1 GTGTTGCCAAGGTGGCGCAACATCCGAGAGAGGTGGGCGAGTGTAACTGCTGTACG 60
DB 828 GTGTTGCCAAGGTGGCGCAACATCCGAGAGAGGTGGGCGAGTGTAACTGCTGTACG 769
QY 61 CCGTTGCTGAACCTCGGTGCGGGCGGCCCCCTGCGCCATCCACACGGCTACTGGAGAAC 120
DB 768 CCGTTGCTGAACCTCGGTGCGGGCGGCCCCCTGCGCCATCCACACGGCTACTGGAGAAC 709
QY 121 GACGGGCGGGCTTCTCTGAGACAGAGCGCGACCGGCTCACCGAACCTTGTGCGCGAACG 180
DB 708 GACGGGCGGGCTTCTCTGAGACAGAGCGCGACCGGCTCACCGAACCTTGTGCGCGAACG 649
QY 181 ACCGTCCTGATGAGCGGCGCTTTCGACTGCTCGATGAGGAGGTGCGGTACCGGACAAACAGCG 240
DB 648 ACCGTCCTGATGAGCGGCGCTTTCGACTGCTCGATGAGGAGGTGCGGTACCGGACAAACAGCG 589
QY 241 CTGCGCGTCGCGCGACAAACCGGATCCAGATCACCGGATCACCGTACCGAGGTCGAA 300
DB 588 CTGCGCGTCGCGCGACAAACCGGATCCAGATCACCGGATCACCGTACCGAGGTCGAA 529
QY 301 GTGGCCATCGCCGCTGATGTCGACGCGAACGCGGACTTAAGCCACCGGGTGAATTTCTCG 360
DB 528 GTGGCCATCGCCGCTGATGTCGACGCGAACGCGGACTTAAGCCACCGGGTGAATTTCTCG 469
QY 361 TGGCGTCGATGCGATGTCCTGCGCTGACCGGACAAATGCTTTGACGCGCGCTGGGCGCATG 420
DB 468 TGGCGTCGATGCGATGTCCTGCGCTGACCGGACAAATGCTTTGACGCGCGCTGGGCGCATG 409
QY 421 CAGTCGCTGTTGAGAGATGTCGAAACCGGACCGTGCATCCGGGAAATCTTTCGAGTACTC 480
DB 408 CAGTCGCTGTTGAGAGATGTCGAAACCGGACCGTGCATCCGGGAAATCTTTCGAGTACTC 349
QY 481 AAACCCGGTGCATCTCTCGGCGCTCACCGAGGTGCTCAAAACGAGAGGCGGCGGCGGATG 540
DB 348 AAACCCGGTGCATCTCTCGGCGCTCACCGAGGTGCTCAAAACGAGAGGCGGCGGCGGATG 289
QY 541 CCGGTGTCGCGGAGACAGGTGCGCGACCGGCGCTTCCGATGCTGCGTGAAGAACTTCTCG 600
DB 288 CCGGTGTCGCGGAGACAGGTGCGCGACCGGCGCTTCCGATGCTGCGTGAAGAACTTCTCG 229
QY 601 GAATGCTGCTGACAGCGGCGGTTTCGAGATCTTCGATTTGGAGAGACGTGCTGCGAGAAC 660
DB 228 GAATGCTGCTGACAGCGGCGGTTTCGAGATCTTCGATTTGGAGAGACGTGCTGCGAGAAC 169
QY 661 CGGTACTTCATGCGCGAGATTCCGCGAAGAGCTGCTGCGCACGACGCGGATCGCGAGC 720
DB 168 CGGTACTTCATGCGCGAGATTCCGCGAAGAGCTGCTGCGCACGACGCGGATCGCGAGC 109
QY 721 AGGTACGGGCGCGCTGTCCCGCGCTGAGGCGCGCGCTGTGCGATTTAAGAAATATGCGC 780
DB 108 AGGTACGGGCGCGCTGTCCCGCGCTGAGGCGCGCGCTGTGCGATTTAAGAAATATGCGC 49
QY 781 CACGACATGGGCTATGCGATTCTGACGCGCGGGAAGCGGTGCGCTGA 828
DB 48 CACGACATGGGCTATGCGATTCTGACGCGCGGGAAGCGGTGCGCTGA 1

RESULT 3
AX089416 50000 bp DNA linear PAT 21-MAR-2001
LOCUS AX089416
DEFINITION Sequence 1 from Patent WO0116303.
ACCESSION AX089416
VERSION AX089416.1 GI:13443677
KEYWORDS

```

SOURCE      Saccharopolyspora spinosa
ORGANISM    Saccharopolyspora spinosa
REFERENCE   1
AUTHORS     Eberz,G., Moenle,V., Froede,R., Velten,R. and Salas,J.A.
TITLE       Nucleic acids which code for the enzyme activities of the spinosyn
JOURNAL     bioynthesis
Patent: WO 0116303-A 1 08-MAR-2001;
BAYER AG (DE)
FEATURES
  source
    1..50000
    /organism="Saccharopolyspora spinosa"
    /mol_type="unassigned DNA"
    /db_xref="taxon:60894"

ORIGIN
Query Match      100.0%; Score 828; DB 6; Length 50000;
Best Local Similarity 100.0%; Pred. No. 9.3e-117;
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGTGTCACAGTGGCCGACCAATGACAGAGTTGGCAATGATGACCTGTACG 60
Db      44797 GGTGTCACAGTGGCCGACCAATGACAGAGTTGGCAATGATGACCTGTACG 44856
QY      61 CCGTTCCTGAATCTCGGTTCGCGGGGCGCCCTCTGCGCCATCCACACGGCTACTGGGGAAC 120
Db      44857 CCGTTCCTGAATCTCGGTTCGCGGGGCGCCCTCTGCGCCATCCACACGGCTACTGGGGAAC 44916
QY      121 GACGGGCGGGCTTCTGTGGAGAGCGCGGACCGGCTCAACGACTTGTGCGCGAAGC 180
Db      44917 GACGGGCGGGCTTCTGTGGAGAGCGCGGACCGGCTCAACGACTTGTGCGCGAAGC 44976
QY      181 ACCGTGCTCGATGCGCGCGCTTTCAGTCTGATGATGAGTGGGTCGGTACCGGACCAACGACG 240
Db      44977 ACCGTGCTCGATGCGCGCGCTTTCAGTCTGATGATGAGTGGGTCGGTACCGGACCAACGACG 45036
QY      241 CTGCGGCTGCGCGGCGGACCAACGGATCCAGATCAACCGGCTCAACCGTCAAGCGAGTGCAC 300
Db      45037 CTGCGGCTGCGCGGCGGACCAACGGATCCAGATCAACCGGCTCAACCGTCAAGCGAGTGCAC 45096
QY      301 GTGGCCATGCGCGCTGATGTCGACGCGGAAACGCGGACTAAGCAGCGGGTGAAGTCTTCTCG 360
Db      45097 GTGGCCATGCGCGCTGATGTCGACGCGGAAACGCGGACTAAGCAGCGGGTGAAGTCTTCTCG 45156
QY      361 TGGGTGATGTCATGTCCTGTCCTGACCCGAGCAATGCTTTTCAGCGCGCGCTGGGCGCATG 420
Db      45157 TGGGTGATGTCATGTCCTGTCCTGACCCGAGCAATGCTTTTCAGCGCGCGCTGGGCGCATG 45216
QY      421 CAGTCCGTGTTGAGATGTCGAAACCGGACCGTTCCTGCGGAAATCTTTCAGATTAATC 480
Db      45217 CAGTCCGTGTTGAGATGTCGAAACCGGACCGTTCCTGCGGAAATCTTTCAGATTAATC 45276
QY      481 AAACCCGCTGAGTCTCTCGGCGGTCAACGAGTGTGTAAGAGAGCGGGCGCGGAGATG 540
Db      45277 AAACCCGCTGAGTCTCTCGGCGGTCAACGAGTGTGTAAGAGAGCGGGCGCGGAGATG 45336
QY      541 CCGGTGTCGCGGAGACAGATGCGGACCGGCGCTTTCGATCTGCTGCTGAGCAATCTTCTG 600
Db      45337 CCGGTGTCGCGGAGACAGATGCGGACCGGCGCTTTCGATCTGCTGCTGAGCAATCTTCTG 45396
QY      601 GAATCGCTGTCGACGCGGGGTTTCAGATCTTTCGATTTGGAGAGACGCTGTGTGAGAGACC 660
Db      45397 GAATCGCTGTCGACGCGGGGTTTCAGATCTTTCGATTTGGAGAGACGCTGTGTGAGAGACC 45456
QY      661 CGGTAATCTTCAATGCGCGAGTTTCGCGAAGAGCTGCTGCGGACGAGCAGGAGATCGCGAC 720
Db      45457 CGGTAATCTTCAATGCGCGAGTTTCGCGAAGAGCTGCTGCGGACGAGCAGGAGATCGCGAC 45516
QY      721 AGGTACGCGGCGCGCTGTTCGCGGCTGGGCGCGCGGCTGTGCGATTATGAGAAATATGCC 780
Db      45517 AGGTACGCGGCGCGCTGTTCGCGGCTGGGCGCGCGGCTGTGCGATTATGAGAAATATGCC 45576

```

```

QY      781 CACGACATGGGCTATGCGATTCGACGCGCGGGAACCGCGTGGCTGA 828
Db      45577 CACGACATGGGCTATGCGATTCGACGCGCGGGAACCGCGTGGCTGA 45624

RESULT 4
LOCUS   AY007564
DEFINITION
  Saccharopolyspora spinosa probable lysR-type transcriptional
  regulator, probable keto acyl reductase, spinosad biosynthetic gene
  cluster, complete sequence, and probable exodeoxyribonuclease V
  genes, complete cds, and unknown gene.
ACCESSION
  AY007564
  AY007564.1 GI:13162633
SOURCE
  ORGANISM
    Saccharopolyspora spinosa
    Saccharopolyspora spinosa
    Bacteria; Actinobacteriia; Actinobacteridae; Actinomycetales;
    Pseudonocardineae; Pseudonocardiaceae; Saccharopolyspora.
  TITLE
    1 (bases 1 to 80161)
    Waldron,C., Matsushima,P., Roetbeck,P.R., Jr., Broughton,M.C.,
    Turner,J., Madduri,K., Crawford,K.P., Merlo,D.J. and Balcz,R.H.
    Cloning and analysis of the spinosad biosynthetic gene cluster of
    Saccharopolyspora spinosa
    JOURNAL
      MEDLINE
      2125765
      11358695
      2 (bases 1 to 80161)
      Waldron,C., Matsushima,P., Roetbeck,P.R., Jr., Broughton,M.C.,
      Turner,J., Madduri,K., Crawford,K.P., Merlo,D.J. and Balcz,R.H.
      Direct Submission
      Submitted (18-AUG-2000) Dow Agrosciences, 9330 Zionsville Rd.,
      Indianapolis, IN 46268 USA
  FEATURES
    source
      1..80161
      /organism="Saccharopolyspora spinosa"
      /mol_type="genomic DNA"
      /db_xref="taxon:60894"
      complement(1135..1971)
      /note="ORF L16"
      /codon_start=1
      /transl_table=11
      /product="probable lysR-type transcriptional regulator"
      /protein_id="AAG23282.1"
      /db_xref="GI:13162654"
      /translation="METRELRYFAVAEELHFGRAAGRLGTAQPLSRITADLEORIG
      VVILQRTSRKSLTEAGAMLTEGRAIIGALAAARRTORAATSPSLVLAAGASG
      ELIALKLDAVPAEPAVAVDILCESQPKTIHDERADVALHQPDPYAEIDITLN
      TEOOVAIILPTSHPLASEPHRMADVSLPDLPLAMPGDGVYPDGPGVEVNNQTLF
      QMIALGRTTVNPPSSRVMLLEGAAVPLDAPDVTTVIAMPFRSRLAGLVNRYAT
      LV"
      2024..2791
      /note="ORF L15"
      /codon_start=1
      /transl_table=11
      /product="probable keto acyl reductase"
      /protein_id="AAG23281.1"
      /db_xref="GI:13162653"
      /translation="MVLVPRRPRTATLESMEQITIALVTGANKGIGYEIAAGIGALGWS
      VGIGARDHORGDVAKLADADGDAFAVSLDTDDASVAAAALLSEERAGRDIVVNN
      AGIAGAWPEEPSTVTPASRAVENNVGIVVTNMLPLRBSRPRVIVNOSHVAS
      IRLQTPRGVDLGIGAYSPSKTFINATITIOYAKLSIDNININNAACPEVATDLNGF
      HGTSTPADGARAIIRLATTPDGPPTGAGFDDAGNVFW"
      complement(3416..4165)
      /gene="spns"
      /gene="spns"
      complement(3416..4165)
      /note="probably involved in fosfamine biosynthesis"
      /gene="spns"
      /codon_start=1
      /transl_table=11
      /evidence="not experimental"
      /product="probable N-dimethyltransferase"
      /protein_id="AAG23280.1"

```

gene /db_xref="GI:13162652"
/translation="MSRVSDFEATSVSVSPDHADIYDAIHSARGRDMAAEGEYVOL
VTRLEPAOSLIDVACGTALERRFAEYAKAGHLEISAPAMEIAIRRPEEPVHIGD
IRDFLGEPEFDVITGLCTFAAAYRTVDLIRVTMRMARHIAAGVAVIEPMFPKFI
DGFVGAVAHGERVIRSLRSHSVLEGRSRMTVRYTVAPETGIRPTEFEIILSTED
ETVALDEDGIAEYLPQAPNGRGLFVGIRN"
/complement (4168. .5325)
/gene="spn"
/complement (4168. .5325)
CDS /gene="spn"
/complement (4168. .5325)
/note="probably involved in forosamine biosynthesis"
/codon_start=1
/evidence=not_experimental
/transl_table=1
/product="probable aminotransferase"
/protein_id="AAG23279.1"
/db_xref="GI:13162651"
/translation="WINTLHQPILGTEBEDATAIEVNASWITGLGPRTRPEAFRHHLG
VDEQVFLNLSGTALFTVQVLDLGPDDVLPISVPAANALASSGARVFCVDV
PRTLNPTLDDVRAITPATKAVLILHYGSPGEVTAIDFCREKGLMIEDSACAVAS
SVHGTACGFGDLATWSPDMKILVYGGMFYADPELAHRARLAVHLEOMGFD
SAKSNRMWDIREDIGORLIGNMTALGSVOLKRLPEINRRREIAOYDRLISDV
BGVLLPRLPGHVSHTFYVQVLAPEIRDOVAQOMLERGITSTRYPPLKVPYIYRA
DCKLSAEDACRKLILPLHPSLDDAEVRVADERQKAVEHHTSORSPLRK"
/complement (5363. .6751)
/gene="spn"
/complement (5363. .6751)
CDS /gene="spn"
/complement (5363. .6751)
/note="probably involved in forosamine biosynthesis"
/codon_start=1
/evidence=not_experimental
/transl_table=1
/product="probable NDP-hexose-3,4-dehydratase"
/protein_id="AAG23278.1"
/db_xref="GI:13162650"
/translation="WQSRKTRALGKGRARVYSCDTCATTEMVVDADRILIASVRYD
HREQSPFVAGSTPIRPSGAVLDEDDVALVEALEIRIAAGNARFESSEFAFEG
LRKALVNSGSANLILASLSLTPGLGARLPGEVITAAVGFPTTINPVOGLVP
VFDVDELGYNATPDIRIKAAVETRAIMLAHGLNPPADEIAEIAKEHEFLVEDN
CDVSGSTRGRLTGTFGDLTVVSPFAHHTSGEGCVLTGSELARLIESLRDGRD
CMCEPVDNCTCKRFPYHIGTLPGYDHKVTSHVGNLKTDTLOALALGOLSKISA
FGSARRRWRRLRELSGLPGILLPVATPHSPSPFGAIIITISADAGTRALVNFLE
SRNIGTLFGSNITRHPAFEQVRYRIADALINSDI VTRTFWGVVPTTDMIDYV
VESIAEFVAKSS"
/complement (7083. .8450)
/gene="spn"
/complement (7083. .8450)
CDS /gene="spn"
/complement (7083. .8450)
/note="involved in forosamine addition"
/codon_start=1
/transl_table=1
/product="probable NDP-forosamyltransferase"
/protein_id="AAG23277.1"
/db_xref="GI:13162649"
/translation="WILGMLPGCSIAIEPMKVLFTPLPASHFENLVPLAMALPAAG
HEVRVAICPNWVMTGAGLTAIPVGDIDLISLAKEMLVSGSVSDEGRHELF
DELSINSGRDITAVEQOLHVDRLIDLMGFAEKMQPDLVWMDAMVSGSVARALG
ARHVMVLVALDYSGLRSGLPEYORSKEPEORVDPDGLWGLAKLAFGATDEELVIG
QATIDIPSMRLPYVDLIDYIMRFPVNGPAVLPTMAERPEPKVCTRLGTRRLS
RVTPEYGRQSDOEOMVERLLRGARLIVETATISDEVEKMELEPNVAVHIVPL
NELRESVLIHHSITTOETRTVNGVQLILPGTFWDESRRAELIADRGAGLVDDPA
LTTEDDVRGOLARLIDEPSFANALIRREIESSPHDIVRLEKLVAREENRTGQ
SDGHP"
/complement (8967. .10427)
/gene="spn"
/complement (8967. .10427)
CDS /gene="spn"
/complement (8967. .10427)
/note="involved in forosamine biosynthesis"
/codon_start=1
/transl_table=1
/product="probable NDP-hexose-2,3-dehydratase"
/protein_id="AAG23276.1"
/db_xref="GI:13162648"

gene /translation="WSSSVEAASAAAPLGSNNTRRVFVLSALSCNGMIPTEFHCWL
ADRGENSEFTENRIJPDRLSKMKDASTENLVHADFPFTEGQVETNNGAASMQ
PIIOAEYGEIIVPEKIDVGLHGMCSKRGVNVNVLQSLPTQOATRSNVAORGS
VPVVDYVIGRGRVIVDYLQSGQSWFYKRRNNVVEOEVPVLPDSCMLTLGQ
VLALLRQDNITNMTRTYLSTPPHDSATGTEBLAASEEPFQAVASISHSIIDSSIS
EAVGMFEBAKARYLRATRVPLSVKMYRTDTSIAHODGXFVAVIASVATREVA
SWTQPMIEPREQGBIALLVKRIIGVLLHVARVAGAKYAEIAEAPVQCSVANYST
PSNDMPPEFLDDITADETVAYESILSEEGCFYQAOQNRRIIEVHEDFAARPPSDFR
MWTIGQGEILRSTHFLNIQARSLVASLHSLMALGR"
/complement (10436. .11434)
/gene="spn"
/complement (10436. .11434)
CDS /gene="spn"
/complement (10436. .11434)
/note="involved in forosamine biosynthesis"
/codon_start=1
/transl_table=1
/product="probable NDP-hexose-3-ketoreductase"
/protein_id="AAG23275.1"
/db_xref="GI:13162647"
/translation="MRKPVRIGVLCASFAWRRLPAMCDVAETEVAVASRDBAKAE
RFARPECEAVLGYORLLERPDIAVVVPLPGKHAEMIGKALBADKHVLAKEPLTTT
ASDPAVLVGLARKNLILRENYLFLHGRHDVVRDILQSGEIGLRREFTVFGIPLP
DTRIRYTELGGALLDTCVYPAAAHFPLIGPLTVYGASHEQESGVDSGVLLQ
SEGTVAHLGYGFVHYRSAYELMSSGRIVVDRAPFPAEWQAVIRERGVDELS
LPAEDQVKAVTARAGTGVDDPAVADSDGSMIQQALVVAIGQARRCGST"
/complement (11530. .12492)
/gene="spn"
/complement (11530. .12492)
CDS /gene="spn"
/complement (11530. .12492)
/note="involved in spinosyn aglycone biosynthesis"
/codon_start=1
/transl_table=1
/product="SpnM"
/protein_id="AAG23274.1"
/db_xref="GI:13162646"
/translation="MPNVAAGTUVLPNIPPREPRITTRAVGTHIGSGOVAPEYLAR
TGTEPETELIAVALDRGMVAVITDYBGLGPTGHTTVYVGRAGQHAMLDAABAQRLG
SGTLTDPVGIWGAQOGQASAPFGEHLPTVAPBLRPAQAQNAVPIDLADIHRNDG
VFTQPLVAGAAVYDLPFDLLEAGRTAVDQVEICAPELVTRFLGRLESDIL
DTSELPROPRARALASVAGRNQSPVPLTVHSTODEIVPFAFGRLRDSYAAAT
PVRRHPSGLAPPAALASSRVVYVSMWDEHRSALSGRDAR"
/complement (12696. .13547)
/gene="spn"
/complement (12696. .13547)
CDS /gene="spn"
/complement (12696. .13547)
/note="involved in spinosyn aglycone biosynthesis;
possibly involved in polyketide bridging"
/codon_start=1
/transl_table=1
/product="methyltransferase-like protein"
/protein_id="AAG23273.1"
/db_xref="GI:13162645"
/translation="MESIFDALHAGRLPHHGYVAGVREDAAGAPWSADADQJLDFI
DKALRFGAHLFDLGCNGQOPVVRACASGVRVTGTVNAOHAAATRLANETLAS
LEPDLVGAQULPYPDGFFQAAMAOVSVOQVDDAAALREHARLLEPGRVLVDGITR
VRLPEEYAAVMTGTATHTINSFTALVEAGFEILIEVDLQOTRCWWSWYDEILRLX
DELAGVPAVAVGTQQRVYLDIAKHGPPAQQLIAVAEYRKHPDYARNEESMGFMLL
QARKQOS"
/complement (13592. .14785)
/gene="spn"
/complement (13592. .14785)
CDS /gene="spn"
/complement (13592. .14785)
Query Match 100.0%; Score 828; DB 1; Length 80161;
Best Local Similarity 100.0%; Pred. No. 8,3e-117;
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GTGTTGCCAGGTGGCGCACCAATCGACAGGTTGGCGGAGATGATGACCTGTGCAAG 60
DB 20168 GTGTTGCCAGGTGGCGCACCAATCGACAGGTTGGCGGAGATGATGACCTGTGCAAG 20227
OY 61 CCGTTGCTGAACCTGGTCGGGGGGGGCCCTGGCCGCAATCCACCAAGGCTACTGGGAAC 120
DB 20228 CCGTTGCTGAACCTGGTCGGGGGGGGCCCTGGCCGCAATCCACCAAGGCTACTGGGAAC 20287

QY 121 GACGGGCGGGCTTCTTCTGCGACGAGCGCGCCGACCGGCTCAACGACTTGTGCGCAACCG 180
 DB 20288 GACGGGCGGGCTTCTTCTGCGACGAGCGCGCGGACCGGCTCAACGACTTGTGCGCAACCG 20347
 QY 181 ACCGCTGCTGATGCGGGCGGTTGACCTGCTGATGCGGGGTGGGGTACCGGACCAACGAGCG 240
 DB 20348 ACCGCTGCTGATGCGGGCGGTTGACCTGCTGATGCGGGGTGGGGTACCGGACCAACGAGCG 20407
 QY 241 CTGCGGCTGCGCGCGCGACCAACGCGATCCAGATCAACCGGATCAACCGTCAACGAGTGCA 300
 DB 20408 CTGCGGCTGCGCGCGCGACCAACGCGATCCAGATCAACCGGATCAACCGTCAACGAGTGCA 20467
 QY 301 GTGGCCATGCGCGCTGATGCGGACCGGAAACCGGACTTAAGCAACCGGCTGCACTTCTG 360
 DB 20468 GTGGCCATGCGCGCTGATGCGGACCGGAAACCGGACTTAAGCAACCGGCTGCACTTCTG 20527
 QY 361 TGGGTCGATGCGCATGCTCCGCGCTGACCGGGAACATGCTTGAACGCGCGCTGCGGCGCATG 420
 DB 20528 TGGGTCGATGCGCATGCTCCGCGCTGACCGGGAACATGCTTGAACGCGCGCTGCGGCGCATG 20587
 QY 421 CAGTCGCTGTTGAGATGTCGCAACCGGACCGGATCCGATCCGAGAAATCCTTCAAGTACTC 480
 DB 20588 CAGTCGCTGTTGAGATGTCGCAACCGGACCGGATCCGATCCGAGAAATCCTTCAAGTACTC 20647
 QY 481 AAACCCGGTGGCATCTCTCGGCGTCAACCGAGTGTCAACGAGAACCGGCGCGCGGAGTG 540
 DB 20648 AAACCCGGTGGCATCTCTCGGCGTCAACCGAGTGTCAACGAGAACCGGCGCGCGGAGTG 20707
 QY 541 CCGGTCCTCGGGGACAGGTGGCGGACCGGCTTCGATGCTGCTGCTGAGCAACTTCTG 600
 DB 20708 CCGGTCCTCGGGGACAGGTGGCGGACCGGCTTCGATGCTGCTGCTGAGCAACTTCTG 20767
 QY 601 GAATCGCTGCTGACAGCGGGGTTGAGATCTCTGATTTGGAGAGACGTTGTCGAGAGAC 660
 DB 20768 GAATCGCTGCTGACAGCGGGGTTGAGATCTCTGATTTGGAGAGACGTTGTCGAGAGAC 20827
 QY 661 CCGTACTTCATGCGCGAGTTGCGCGAAGAGCTGCTGCGCACGACGCGGATTCGGGAC 720
 DB 20828 CCGTACTTCATGCGCGAGTTGCGCGAAGAGCTGCTGCGCACGACGCGGATTCGGGAC 20887
 QY 721 AGGTACGGGCGGCTGTGCGCGGCTGGGCGCGCGGCTGCGGATTAAGAAATATGCC 780
 DB 20888 AGGTACGGGCGGCTGTGCGCGGCTGGGCGCGCGGCTGCGGATTAAGAAATATGCC 20947
 QY 781 CACGACATGGGCTATGCGATTCGACGCGCGGAAAGCGGCTGCGCTGA 828
 DB 20948 CACGACATGGGCTATGCGATTCGACGCGCGGAAAGCGGCTGCGCTGA 20995

RESULT 5
 ARI65018 80161 bp DNA linear PAT 17-OCT-2001
 LOCUS Sequence 1 from patent US 6274350.
 DEFINITION ARI65018
 ACCESSION ARI65018
 VERSION ARI65018.1 GI:16238344
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 Unclassified.

REFERENCE 1 (bases 1 to 80161)
 AUTHORS Baitz,R.H., Crawford,K.P., Broughton,M,Christine., Madduri,K.,
 Merlo,D.J., Turner,J.R., Treadway,P.J. and Waldron,C.
 TITLE Biosynthetic genes for spinosyn insecticide production
 JOURNAL Patent: US 6274350-A 14-AUG-2001;
 FEATURES Location/Qualifiers
 source 1..80161
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN
 Query Match 100.0%; Score 828; DB 6; Length 80161;
 Best Local Similarity 100.0%; Pred. No. 8.3e-117;

Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGTTGCCAGGTGGGCGACCAACATCGACGAGTTGGGAGATGATGACCTGCTGACG 60
 DB 20168 GTGTTGCCAGGTGGGCGACCAACATCGACGAGTTGGGAGATGATGACCTGCTGACG 20227
 QY 61 CCGTTGCTGAACCTCGGTGCGGGGCGGCGCTGCGGCATCAACGAGCTACTGAGAGAAC 120
 DB 20228 CCGTTGCTGAACCTCGGTGCGGGGCGGCGCTGCGGCATCAACGAGCTACTGAGAGAAC 20287
 QY 121 GACGGGCGGGCTTCTTCTGCGACGAGCGCGCGGATCCAGATCAACCGTTCGCGCAACG 180
 DB 20288 GACGGGCGGGCTTCTTCTGCGACGAGCGCGCGGATCCAGATCAACCGTTCGCGCAACG 20347
 QY 181 ACCGCTGCTGATGCGGGCGGTTGACCTGCTGATGCGGGGTGGGGTACCGGACCAACGAGCG 240
 DB 20348 ACCGCTGCTGATGCGGGCGGTTGACCTGCTGATGCGGGGTGGGGTACCGGACCAACGAGCG 20407
 QY 241 CTGCGGCTGCGCGCGCGACCAACGCGATCCAGATCAACCGGATCAACCGTCAACGAGTGCA 300
 DB 20408 CTGCGGCTGCGCGCGCGACCAACGCGATCCAGATCAACCGGATCAACCGTCAACGAGTGCA 20467
 QY 301 GTGGCCATGCGCGCTGATGCGGACCGGAAACCGGACTTAAGCAACCGGCTGCACTTCTG 360
 DB 20468 GTGGCCATGCGCGCTGATGCGGACCGGAAACCGGACTTAAGCAACCGGCTGCACTTCTG 20527
 QY 361 TGGGTCGATGCGCATGCTCCGCGCTGACCGGGAACATGCTTGAACGCGCGCTGCGGCGCATG 420
 DB 20528 TGGGTCGATGCGCATGCTCCGCGCTGACCGGGAACATGCTTGAACGCGCGCTGCGGCGCATG 20587
 QY 421 CAGTCGCTGTTGAGATGTCGCAACCGGACCGGATCCGATCCGAGAAATCCTTCAAGTACTC 480
 DB 20588 CAGTCGCTGTTGAGATGTCGCAACCGGACCGGATCCGATCCGAGAAATCCTTCAAGTACTC 20647
 QY 481 AAACCCGGTGGCATCTCTCGGCGTCAACCGAGTGTCAACGAGAACCGGCGCGGAGTG 540
 DB 20648 AAACCCGGTGGCATCTCTCGGCGTCAACCGAGTGTCAACGAGAACCGGCGCGGAGTG 20707
 QY 541 CCGGTCCTCGGGGACAGGTGGCGGACCGGCTTCGATGCTGCTGCTGAGCAACTTCTG 600
 DB 20708 CCGGTCCTCGGGGACAGGTGGCGGACCGGCTTCGATGCTGCTGCTGAGCAACTTCTG 20767
 QY 601 GAATCGCTGCTGACAGCGGGGTTGAGATCTCTGATTTGGAGAGACGTTGTCGAGAGAC 660
 DB 20768 GAATCGCTGCTGACAGCGGGGTTGAGATCTCTGATTTGGAGAGACGTTGTCGAGAGAC 20827
 QY 661 CCGTACTTCATGCGCGAGTTGCGCGAAGAGCTGCTGCGCACGACGCGGATTCGGGAC 720
 DB 20828 CCGTACTTCATGCGCGAGTTGCGCGAAGAGCTGCTGCGCACGACGCGGATTCGGGAC 20887
 QY 721 AGGTACGGGCGGCTGTGCGCGGCTGGGCGCGCGGCTGCGGATTAAGAAATATGCC 780
 DB 20888 AGGTACGGGCGGCTGTGCGCGGCTGGGCGCGCGGCTGCGGATTAAGAAATATGCC 20947
 QY 781 CACGACATGGGCTATGCGATTCGACGCGCGGAAAGCGGCTGCGCTGA 828
 DB 20948 CACGACATGGGCTATGCGATTCGACGCGCGGAAAGCGGCTGCGCTGA 20995

RESULT 6
 BD137649 80161 bp DNA linear PAT 18-SEP-2002
 LOCUS BD137649
 DEFINITION Biosynthetic genes for spinosyn insecticide production.
 ACCESSION BD137649
 VERSION BD137649.1 GI:23232594
 KEYWORDS JP 2002505881-A/1.
 SOURCE Saccharopolyspora spinosa
 ORGANISM Saccharopolyspora spinosa
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Pseudonocardiales; Pseudonocardaceae; Saccharopolyspora.

REFERENCE 1 (bases 1 to 80161)
 AUTHORS Baitz,R.H., Broughton,M.C., Crawford,K.P., Madduri,K., Merlo,D.J.,
 Treadway,P.J., Turner,J.R. and Waldron,C.


```

Db      20648 AACCCTGCGCATCTCTGCGCGTCAACCGAGTCTGTAACGAGAACGCGCGCGCGGATG 20707
Oy      541 CCGGTGTCCGGGAGACAGTGGCCGACCGGCTTCGGATCTGCTGGCTGAGCACTTCTG 600
Db      20708 CCGGTGTCCGGGAGACAGTGGCCGACCGGCTTCGGATCTGCTGGCTGAGCACTTCTG 20767
Oy      601 GAATGCTGCGGCGAGCGCGGGGTTCCAGATCTCTGATTTGGAGAGAGCTGTCTGAGGAGAC 660
Db      20768 GAATGCTGCGGCGAGCGCGGGGTTCCAGATCTCTGATTTGGAGAGAGCTGTCTGAGGAGAC 20827
Oy      661 CCGTACTTATGCTCCGAGATTCGCCAAGAGCTCTGCTGCGCACGACGAGCGGATTCGCGGAC 720
Db      20828 CCGTACTTATGCTCCGAGATTCGCCAAGAGCTCTGCTGCGCACGACGAGCGGATTCGCGGAC 20887
Oy      721 AGGTACGGGCGGCTGTGCGCGCGCTGCGGCGCGCGCTGCTGCGATTTAGAAATATGCC 780
Db      20888 AGGTACGGGCGGCTGTGCGCGCGCTGCGGCGCGCGCTGCTGCGATTTAGAAATATGCC 20947
Oy      781 CACGACATGGGCTATTCGATTTCTGACGCGGCGGAAACCGGTCGGCTGA 828
Db      20948 CACGACATGGGCTATTCGATTTCTGACGCGGCGGAAACCGGTCGGCTGA 20995

RESULT 8
LOCUS   AX600587 36538 bp DNA linear PAT 14-FEB-2003
DEFINITION
Sequence 2 from Patent WO02079477.
ACCESSION
AX600587
VERSION
AX600587.1 GI:28400303
KEYWORDS
SOURCE
Saccharopolyspora sp. NRRL 30141
ORGANISM
Saccharopolyspora sp. NRRL 30141
Bacteria; Actinobacteriales; Actinomycetales;
Pseudonocardiales; Pseudonocardiales; Saccharopolyspora.
REFERENCE
1 Hahn, D. R., Jackson, J. D., Bullard, B. S., Gustafson, G. D., Waldron, C.
and Mitchell, J. C.
Biothetic genes for butenyl-epinosyn insecticide production
Patent: WO 02079477-A 2 10-OCT-2002;
Dow Agrosciences LLC (US)
FEATURES
Location/Qualifiers
1..36538
/organism="Saccharopolyspora sp. NRRL 30141"
/mol_type="unassigned DNA"
/db_xref="taxon:221471"
/note="LW107129"
ORIGIN
Query Match 90.5%; Score 749.6; DB 6; Length 36538;
Best Local Similarity 94.1%; Pred. No. 8e-105;
Matches 779; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Oy      1 GGTGGCAGGTGGCGGACCAACATCGACAGGTTGGGAGATGATGACTGTGACG 60
Db      938 GTGTGGCAGGTGGCGGACCAACATCGACAGGTTGGGAGATGATGACTGTGACG 879
Oy      61 CCGTGTGTAATCTCGGTGCGGCGGCGGCTTCGCCATTCACCAAGGCTTACGGAAGAC 120
Db      878 CCGTGTGTAATCTCGGTGCGGCGGCGGCTTCGCCATTCACCAAGGCTTACGGAAGAC 819
Oy      121 GACGGGCGGCTTCTGTGGCAGCGCGCGGCTTCGCCATTCACCAAGGCTTCTGTGGCAGCG 180
Db      818 GACGGGCGGCTTCTGTGGCAGCGCGCGGCTTCGCCATTCACCAAGGCTTCTGTGGCAGCG 759
Oy      181 ACCGTGCTGATGGCGGCTTCTGTGGCAGCGCGCGGCTTCGCCATTCACCAAGGCTTCTGTGGCAGCG 240
Db      758 ACCGTGCTGATGGCGGCTTCTGTGGCAGCGCGCGGCTTCGCCATTCACCAAGGCTTCTGTGGCAGCG 699
Oy      241 CTGCGGCTGCGCGCGGAGCAACGCGATTCAGATTCACCGGCTTCTGTGGCAGCGGAGCA 300
Db      698 CTGCGGCTGCGCGCGGAGCAACGCGATTCAGATTCACCGGCTTCTGTGGCAGCGGAGCA 639
Oy      301 GTGGCCATCGCGGCTTCTGTGGCAGCGGAGCAACGCGATTCAGATTCACCGGCTTCTGTGGCAGCG 360

```

```

Db      638 GGTGCAATCGCGGTTGATTGGGCGCGGACCGCGGACTAAGCACAGGAGACTTCTCG 579
Oy      361 TCGGTGATGCCATTCCTCTGCGGCTTCGATCCCGGACCAATGCTTTGAGCCGCGCGGCGCATG 420
Db      578 TCGGTGATGCCATTCCTCTGCGGCTTCGATCCCGGACCAATGCTTTGAGCCGCGGCGCATG 519
Oy      421 CAGTGCCTGTTGGAATGTCGGAACCGGACCGGTCATTCGGGAAATCTTCTGATGACTC 480
Db      518 CAGTGCCTGTTGGAATGTCGGAACCGGACCGGTCATTCGGGAAATCTTCTGATGACTC 459
Oy      481 AACCCTGCGATCTCGCGCTCAACGAGTCTCAACGAGAACGCGCGCGGATG 540
Db      458 AACCCTGCGATCTCGCGCTCAACGAGTCTCAACGAGAACGCGCGGATG 399
Oy      541 CCGGTGTCCGGGAGACAGTGGCGGACCGGCTTCGGATCTGCTGCGGCTGAGCACTTCTG 600
Db      398 CCGGTGTCCGGGAGACAGTGGCGGACCGGCTTCGGATCTGCTGCGGCTGAGCACTTCTG 339
Oy      601 GAATGCTGCGGCGGCGGCTTCGAGATCTCTGATTTGGAGAGAGCTGTCTGAGAGAC 660
Db      338 GAATGCTGCGGCGGCGGCTTCGAGATCTCTGATTTGGAGAGAGCTGTCTGAGAGAC 279
Oy      661 CCGTACTTATGCTCCGAGATTCGCCAAGAGCTCTGCTGCGCACGACGAGATTCGCGGAC 720
Db      278 CCGTACTTATGCTCCGAGATTCGCCAAGAGCTCTGCTGCGCACGACGAGATTCGCGGAC 219
Oy      721 AGGTACGGGCGGCTGTGCGCGGCTGCGGCGCGGCTGCTGCGATTTAGAAATATGCC 780
Db      218 AGGTACGGGCGGCTGTGCGCGGCTGCGGCGCGGCTGCTGCGATTTAGAAATATGCC 159
Oy      781 CACGACATGGGCTATTCGATTTCTGACGCGGCGGAAACCGGTCGGCTGA 828
Db      158 GACGACATGGGCTATTCGATTTCTGACGCGGCGGAAACCGGTCGGCTGA 111

```

```

RESULT 9
LOCUS   AX598613 846 bp DNA linear PAT 14-FEB-2003
DEFINITION
Sequence 21 from Patent WO02088176.
ACCESSION
AX598613
VERSION
AX598613.1 GI:28398761
KEYWORDS
SOURCE
Streptomyces platensis
ORGANISM
Streptomyces platensis
Bacteria; Actinobacteriales; Actinomycetales;
Streptomycinales; Streptomycetales; Streptomyces.
REFERENCE
1 Farnet, C. M., Zazopoulos, E., Staffa, A. and Yang, X.
Genes and proteins for the biosynthesis of polyketides
Patent: WO 02088176-A 21 07-NOV-2002;
Ecopia Biosciences Inc. (CA)
FEATURES
Location/Qualifiers
1..846
/organism="Streptomyces platensis"
/mol_type="unassigned DNA"
/db_xref="taxon:58346"
/note="Subsp. robaceus"
ORIGIN
Query Match 20.5%; Score 170; DB 6; Length 846;
Best Local Similarity 59.4%; Pred. No. 1.9e-16;
Matches 348; Conservative 0; Mismatches 215; Indels 9; Gaps 2;

Oy      5 TGCAGAGTGGCGGACCAACATCGACAGGTTGGGAGATGATGACTGTGACGCGT 64
Db      35 TCCCGGTCTCCGCGCGCGCTCCCGAAGAGTGGACACTTACGACCGCTTCACCGCAC 94
Oy      65 TGTGTAATCTCGGTGCGGCGGCGGCTTCGCCATTCACCAAGGCTTCTGTGGCAGCG 124
Db      95 TGTGTAATCTCGGTGCGGCGGCGGCTTCGCCATTCACCAAGGCTTCTGTGGCAGCG 148
Oy      125 GCGGCGCTTCTGTGGCAGCGGAGCAACGCGGCTTCTGTGGCAGCG 181

```

149 ACAACGACACCCCGCTTCGTGAAAGCGCCGACCGGCTCACCGACACGATGACCGACCGCC 208
182 CCGTGTCTGATATGGACGGCGCTTGCAGTCTGATATGTGGAGGTCGGTATCCGGACAAACAGACGCG 241
209 TGGGATATGACCAAGGAGACAGCGGGTCTCTGACGTGCGGTGTCGGAGATCGGCCAGCGCGCCA 268
242 TGGCGGTGCGCGCGCACAAACCGGATTCAGATACCGGACATCAACGTCAGCCAGTGGTCAAG 301
269 TGGGATATGCGCCGGGCGACACCGCGGCCATATGTCACCGGAGATCGCGATCAGCAAGACAGAA 328
302 TGGCCATATGCGCGCTTATATTCGCGACCGCGAACGCGGACTAAGCCACCGGATGGAATTCTGCT 361
329 TCGCCCGGCGCACCCCGCTCGCGAGGCGCGCGCTGAGCCACCGCGTGGAGTTCCGCGC 388
362 GCGTGGATATGCATATGTCCTGCGCGCTATCCCGGACAATATGTTTCGACGCGCGCTGGGGCATGCG 421
389 ACGCGGACGCAATGGAACTGCGCTTCTCCCGAGACTCTTTCAGACGCGCGCATCGCCATGCG 448
422 AGTCGCTTTGGAGATATGTCGGAACCGGACCGTGCATTCGGGAAATCTTTCGAGTACTCA 481
449 AGTGGATCTTCCACATATGCCCGACCGCGAGCGGGTCTCGCGCAGATCGGCCCGCTATCTGC 508
482 AACCCGATGGCATCTCTCGCGCTACCGAGTGTCAACAGAAACGCGCGCGCGGAGATGC 541
509 GCCCGGCGCGCGCTGGTCTTCAACGATTTTCGAGCGCGCGCCCGTCCCCCGCAGAA 568
542 CGGTATCCGGGG 553
569 AGCAACCGCGCG 580

RESULT	10				
LOCUS	AX598593				
DEFINITION	Sequence 1 from Patent WO0208176.	52101 bp	DNA	linear	PAT 14-FEB-2003
ACCESSION	AX598593				
VERSION	AX598593.1	GI:28398751			
KEYWORDS					
SOURCE					
ORGANISM	Streptomyces platensis				
	Streptomyces platensis				
	Bacteria; Actinobacteria; Actinomycetales;				
	Streptomycineae; Streptomycetaceae; Streptomyces.				
REFERENCE					
	1 Farnet, C.M., Zazopoulos, E., Staffe, A. and Yang, X.				
	Genes and proteins for the biosynthesis of polyketides				
	Patent: WO 0208176-A 1 07-NOV-2002;				
	Ecopia Biosciences Inc. (CA)				
FEATURES					
	location/Qualifiers				
	1..52101				
	/organism="Streptomyces platensis"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:58346"				
	/note="subsp. rosaceus"				

Query Match	Similarity	20.5%	Score 170	DB 6	Length 52101
Best Local	Similarity	59.4%	Pred. No. 7.4e-17		
Matches	328	Conservative	0	Mismatches 215	Indels 9
				Gaps	2
QY	5	TGCCAGGTGGGGCCCAACAATCGACAGAGTTGGGCGAGATGTATGACTGGTACGCGCT	64		
Db	51268	TCCCCGTTCGCGCCCCCGTCCCGAAGAGGTTCGACACTTACGACCGCGCTACCGCAC	51327		
QY	65	TGCTGAATCGGTTCGGGGGGGGCCCCCTTGGCCATTCACACGGGCTACTGGGAAACGACG	124		
Db	51328	TGGAACACGGAAGCGCGCGCG-----GCAAGCTTCACTTCGGCTACTGGGAGCTGACG	51381		
QY	125	GGCGGGTTCCTCTGGC---AGCAGGCGCGCGACCGGGCTACACGACCTTGTCCGCGGAACGGA	181		
Db	51382	ACAAACGACACCCGCTGTGTGGAAACGGCGCGACGGGCTCAACCGGACACGATGACCGACCGCC	51441		
QY	182	CCGTGTCATGAGCGGCGTTTCGACTGTCTCATGTGGGGTGGCGGTACCGGACCAACGACGC	241		

Db 51442 TGCAGATCGACCAAGGGACAGCGGATCCTTCGACCTCGGCTGCGGAGTGGCCAGCGGCCCA 51501

Qy 242 TGCGCGTCGCGCGCGACAAACGCGATCGAATGCAACCGGCATCAACCGTCAGCCAGTCGACAG 301

Db 51502 TGCAGATCGCCCGCGCCGACCGCGGCCCAATGTCAACGGGCATTCGCGATCAACAAGAACACA 51561

Qy 302 TGCGCATCGCGCGCGTGATTTGCGCAACGCGAAACGCGAATTAAGCCACCGGATGAGATTCTCGT 361

Db 51562 TCGCCCGCGCCACCGCCCTTCGCGGAGGCGCGGCTTGAGCGACCGCGCTGAGATTTCGCGC 51621

Qy 362 GCGTCGATGCGCATGTCCCTGCGCGTACCCCGAACAAATCTTTGCAACGCGCGCTGAGCCATGCG 421

Db 51622 ACGCGCAAGCGCATGGAATCTGCTTCCCGACGACTCTTTGCAACGCGCGCATTCGCATCG 51681

Qy 422 AGTCGCTGTGGAGAGTGTCCGAACCGGACCGGTCCATTCGGGGAAATCTTTCGAGTACTCA 481

Db 51682 AGTCGATCTTCCAAATGCCGACCGGAGAGGGATCTTCGCGGAGATCCGCCCGTACTTCG 51741

Qy 482 AACCCGCTGCGATCTCTGCGGCTCAACCGAGGTGCTCAAAAGAAAGAGGGCGCGGGAGTGC 541

Db 51742 GCCCGCGCGCGCGCTGTGTCCTCACCGACTTCTTTGAGGCGCGGCCCGTCCCGCCGAGA 51801

Qy 542 CGGTGTCGCGGG 553

Db 51802 AGCAGCGCGCGG 51813

```

RESUT 11
AX089434
LOCUS AX089434 900 bp DNA linear PAT 21-MAR-2001
DEFINITION Sequence 19 from Patent WO0116303.
ACCESSION AX089434
VERSION AX089434.1 GI:13443695
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
    source
        location/Qualifiers
            1..900
            /organism="Saccharopolyspora spinosa"
            /mol_type="unassigned DNA"
            /db_xref="taxon:60894"
            1..900

```

		20.0%;	Score 165.4;	DB 6;	Length 900;	
	Query Match					
	Best Local Similarity	52.8%;	Pred. No. 9.2e-16;			
	Matches 386;	Conservative % 0;	Mismatches 336;	Indels	9;	Gaps 1;
QY	15	CGACCAACATCGAGCAGGTTGGCGAGATGATGACCTGGTCAAGCGCGTTGTGAATCTC	74			
Db	21	CACACAGGCAATTTTCGACCGCGGATCAGGTGAGATTCATCTTCGATGATGGCGGACAGG	80			
QY	75	GGTCGCGGGCGGCCCCCTCGCGCATTCACACACGCGTACTCTGGAGAAACGACGCGGCGGGCTTC	134			

[illegible]

Query Match	20.0%: Score 165.4; DB 6; Length 6085;
Best local similarity 60.6%; Pred. No. 5.9e-16;	
Matches 271; Conservative 0; Mismatches 176; Indels 0; Gaps 0;	
<p>ORIGIN</p> <p>Query Match</p> <p>Best local similarity 60.6%; Pred. No. 5.9e-16;</p> <p>Matches 271; Conservative 0; Mismatches 176; Indels 0; Gaps 0;</p>	<p>/function="ORF"</p> <p>4013..4999</p> <p>/function="ORF"</p> <p>m1sc_RNA</p> <p>5071..6085</p> <p>/function="ORF"</p>
<p>90 CTGCGCCATCCACCAACCGCTACTGTGGAGAAAGACGAGCGGGCTTCTTGCGACAGCGCCG 149</p> <p>1833 CGGCAACATCCACGCTGGGGTACTGGGAGCAGACCCCGAGGTGCCGATCGCCGAGGCCAC 1892</p> <p>150 CGACCGGCTCACGCACTTGTGCGCGAAGCAGACCGTGTCTGATGGCGGCGTTGCACTGCT 209</p> <p>1893 CGACCGGCTCACCGATCTCGTCGCGCGAGCGCTTCGCGCTCGCCCGACCGGCATCTGCT 1952</p> <p>210 CGATGTTGGGGTGGCGGTACCGGAGCAACAGAGCGTGTGCGCGCGACCAACGCGATTCGA 269</p> <p>1953 GGACGTGGGCTGGCGGATCGGCGTGGCGGCTCTTCAAGATCGCGGAGCGACGACGCTCG 2012</p> <p>270 GATCACCGGCTCACCGCTACGCGACGAGTGCAGTGGCCATCGCGCTGATTTGGCGACCGGA 329</p> <p>2013 CGTCAACCGGATTCACCGTCAAGCGACGACGAGGTCAACGAGCGGCGCGGCGGTTGGA 2072</p> <p>330 ACGCGGACTAAGCCACCGGGTGGACTTCTGTGCGTGCATATGTCATTCCTTGCCTGACC 389</p> <p>2073 GTCCGATGTCGCGGGGCGGGGCTCTCTTCCGCTGGCGGACGCCATGACCTTCCCTTGA 2132</p> <p>390 GGACAAATGCTTTCAGACGCGCGCTGGGGCCATGCGCTGTTGGAGATGTCCGAACCGGA 449</p> <p>2133 GAGCGTCTCTTCAGACGCGCGCTTCCGCAATGATGCTGCTGCATCTGCCGACACGAC 2192</p> <p>450 CCGTGCCATCCGGGAAATCTTTCAGATCTCAAAACCGGTGGCATCTCGGCGTCAACGA 509</p> <p>2193 ACCGCGCTCAAGAGATTCACCGGGTGTGCGGCGCGGCGGCGGCTGTCATGCGCGA 2252</p> <p>510 GGTGCTCAACGAGAAAGCGGCGGCGG 536</p> <p>2253 CCTGTGTACGACGACGCGCTTCAACCGG 2279</p>	<p>RESULT 13</p> <p>ARI44763 6085 bp DNA linear PAT 08-AUG-2001</p> <p>LOCUS Sequence 4 from patent US 6210935.</p> <p>DEFINITION</p> <p>ACCESSION ARI44763</p> <p>VERSION ARI44763.1 GI:15106630</p> <p>KEYWORDS</p> <p>SOURCE Unknown.</p> <p>ORGANISM Unknown.</p> <p>REFERENCE 1 (bases 1 to 6085)</p> <p>AUTHORS Schnupp,T., Engel,N., Bietenhader,J., Toupet,C. and Pospiech,A.</p> <p>TITLE Staurosporin biosynthesis gene clusters</p> <p>JOURNAL Patent: US 6210935-A 4 03-APR-2001;</p> <p>FEATURES</p> <p>source 1..6085</p> <p>/organism="unknown"</p> <p>/mol_type="unasigned DNA"</p>
<p>150 CGACCGGCTCACCGCTTGTGCGCGAAGCAGCGGTGCTCATGCGGCGCTTGCATCTCT 209</p>	

Db 1893 CGACCGGCTCACCGATCTGTGCGCCGAGCGCCTCGGCTCCGCCCGACCGGCATCTGCT 1952
Qy 210 CGATGCGGGGCTGGGATCCGGAACCGACGCTGCGCGCTCGGCGCGCAACAGCATTCGA 269
Db 1953 GGAAGTGGGCTCGGGATCGGCTGCGCGCTTCAGATCGCGGAGCGCGACGATCGC 2012
Qy 270 GATCAACCGGCATCAACGCTGACCGAGTGCAGTGCATCGCGCTGATTGCGACGCA 329
Db 2013 CGTCAACCGGATCAACGCTGACCGAGTGCAGTGCAGGCGCGCGAGCGGCGGTGA 2072
Qy 330 AGCGGACCTAACCGACCGGGTGAATCTTCGTCGCTGCATGCTCCCTCGCTACCC 389
Db 2073 GTCCATGCGCGGGCGCGGCTCTTCGCTGCGGAGCGCATGACCTCCCTTGA 2132
Qy 390 GGACATGCTTTCAGACGCGCGCTGGGCAATGAGTGCCTTGGAGATGCGCAACCGGA 449
Db 2133 GAACTCTCTTCAGACGCGCGCTTGCCTTCAGATGCTGCTGATCTGCCGACGAC 2192
Qy 450 CGTGCATCCGGAATCTTCAGTACTCAACCGCGTGCATCTCGCGTACCGA 509
Db 2193 ACCCGGCTCAAGAGATCAACCGGGTGTCCGCCCGCGCGCGCTGTGATCGCGA 2252
Qy 510 GGTCTCAACGAGAACCGCGCGCGG 536
Db 2253 CTTGTGTCGCGACGCGTTCACCGG 2279

RESULT 14
AX089419
LOCUS AX089419 45624 bp DNA linear PAT 21-MAR-2001
DEFINITION Sequence 4 from Patent WO0116303.
ACCESSION AX089419
VERSION AX089419.1 GI:13443680
KEYWORDS
SOURCE
ORGANISM
Saccharopolyspora spinosa
Saccharopolyspora spinosa
Bacteria; Actinobacteria; Actinomycetales;
Pseudonocardiales; Pseudonocardiales; Saccharopolyspora.

REFERENCE
AUTHORS Eberz, G., Moehrle, V., Froede, R., Velten, R. and Salas, J.A.
TITLE Nucleic acids which code for the enzyme activities of the spinosyn biosynthesis
JOURNAL Patent: WO 0116303-A 4 08-MAR-2001;
BAYER AG (DE)
FEATURES
source location/Qualifiers
1..45624
/organism="Saccharopolyspora spinosa"
/mol_type="unassigned DNA"
/db_xref="taxon:60894"

ORIGIN

Query Match 20.0%; Score 165.4; DB 6; Length 45624;
Best Local Similarity 52.8%; Pred. No. 3.7e-16;
Matches 386; Conservative 0; Mismatches 336; Indels 9; Gaps 1;

Qy 15 CGCAACAATCGACAGATTGGGAGATGATGACCTGGTACGCGTTGCTGAATC 74
Db 7421 CACACAGGATTTCCGACCGGAGATCAGTGAATCTTCATGCTTGGCGCAAG 7480
Qy 75 GGTGCGGGCGGCGCTTGGCCATCACACGCGTACTGGAGAAAGACGCGGCGCTTC 134
Db 7481 GCGTCCCTGACACAGGTTACTGGCGCGGATGCGGAGGATGCGGTCACAC 7540
Qy 135 CTGGCAGAGCGCGCGACCGGCTGACCGACCTTGTCCCGCAACGAGCGGTGCAAG 194
Db 7541 GTGTGCGATGTCGCGACCACTGACCGACCTTTCATCAACAGCGCGCTCGTCC 7600
Qy 195 CGGCGTGAATGCTGATGTTGGGCTGCGTACCGCAACAGCGCTGCGCTGCGCG 254
Db 7601 CGGAGCGACCTGTTGACCTGGGCTGGGCAATGGGAGCGCCGTAATGTCGTTGGGCAATG 7660
Qy 255 CGACAGCGGATCCAGATCAACGCGATACCGTCAACGAGGTGCAATGTCGATCGCGC 314

Db 7661 CGCAGCGCGCTTCAGATCACCGGAATACCGTGAAGCCCGAGATCTCGCGCGCGAC 7720
Qy 315 TGATTGCGGACCGGAACCGGATCAAGCACCGGATGATCTTCGTCGCTGATGCAAT 374
Db 7721 CAGGCTTCGCAACGAGACCGGATCGCGGAGTCTTGAATGATCTTATGACGCGGC 7780
Qy 375 GTCCCTGCGGATACCGGACCAATGCTTTCAGACCGCGCTGGGCAATGCAATGCTGTTGA 434
Db 7781 CCAAGTGCCTTACCGGACGCTTCTTCAAGCCGCGCATGGGCGATGCAATCGCTGTCGA 7840
Qy 435 GATGTCGAACCGGACCGTGCATCCGGGAAATCTTTCGATGACTCAAAACCGGTGCAAT 494
Db 7841 GATGTCGAACCGGACCGGATCCGGAATCCGGAATCCGGAATCCGCGCGCGG 7900
Qy 495 CCGGCGCTTACCGAGTCTGCAACGAGAACCGGCGCGCGGATCCGCTGTCGCGGA 554
Db 7901 GTTCTCTTCGAGACATCATCTCGGCTTTCGATCTCCCGAAGATCCGCGGCTTTC 7960
Qy 555 CAGTGGCGGACCGGCTTCGATCTGCTGCTGAGCACTTTCGAAATGCTGCTGTC 614
Db 7961 GACGGGACGACCGGCTTACCTT-----GACAGCTTCAAGCGCTGCTGACGCA 8011
Qy 615 AGCGGCTTCAGATCTCTGATGGAAGACGTCGTCGAGAACCGGTACTTCATGCC 674
Db 8012 AGCGGCTTCAGATCTCTGAAATGACCGACCTCACGACGACAGACGATGATGCTTC 8071
Qy 675 GCAGTGGCGGACGCTGCTGCGCACGACGAGATGCGGACGATGCGGCGCGC 734
Db 8072 CTGTGATGTCAGAGATGCTCCGAAATCTGATGAGCTGCGGCTGAGCTGCGCGC 8131
Qy 735 TGTGCGCGGCT 745
Db 8132 TGTGCGGACCT 8142

RESULT 15
AX089416/C
LOCUS AX089416 50000 bp DNA linear PAT 21-MAR-2001
DEFINITION Sequence 1 from Patent WO0116303.
ACCESSION AX089416
VERSION AX089416.1 GI:13443677
KEYWORDS
SOURCE
ORGANISM
Saccharopolyspora spinosa
Saccharopolyspora spinosa
Bacteria; Actinobacteria; Actinomycetales;
Pseudonocardiales; Pseudonocardiales; Saccharopolyspora.

REFERENCE
AUTHORS Eberz, G., Moehrle, V., Froede, R., Velten, R. and Salas, J.A.
TITLE Nucleic acids which code for the enzyme activities of the spinosyn biosynthesis
JOURNAL Patent: WO 0116303-A 1 08-MAR-2001;
BAYER AG (DE)
FEATURES
source location/Qualifiers
1..50000
/organism="Saccharopolyspora spinosa"
/mol_type="unassigned DNA"
/db_xref="taxon:60894"

ORIGIN

Query Match 20.0%; Score 165.4; DB 6; Length 50000;
Best Local Similarity 52.8%; Pred. No. 3.6e-16;
Matches 386; Conservative 0; Mismatches 336; Indels 9; Gaps 1;

Qy 15 CGCAACAATCGACAGATTGGGAGATGATGACCTGTCACGCGTTGCTGAATC 74
Db 38204 CACACAGGATTTCCGACCGGATGAGTGAATCTTCATGCTTGGCGCAAG 38145
Qy 75 GGTGCGGGCGGCGCTTGGCCATCACACGCGTACTGGAGAAAGACGCGGCGCTTC 134
Db 38144 GCGTCCCTGACACAGGTTACTGGGCGGCGGATTCGGGAGATCCGCTGACAC 38085
Qy 135 CTGGCAGAGCGCGCGACCGGCTGACCGACCTTGTGCGGAAAGACGCTGCTGATG 194

Db 38084 GTGTCGATGCTGCGACCAACTGACCACTCTTCATCGACAAAGCCGCGCTCCGTCC 38025
QY 195 CCGCGTTGACTGCTCGATGAGGGTGCAGTACCGGACAAACGCGCTGCGCGTCCGCGC 254
Db 38024 CCGAGCGGACCTGTTGACCTGGGGTGGCGCAATGGGCAAGCCGTAAGTCCGTGCGGCATG 37965
QY 255 CGACAAACCGGATCCAGATACCGGGCATCACCGTCAAGCCAGTGCATAGTGGCCATCGCCGC 314
Db 37964 CCGCAGCGGCGTTGCAAGTACCGGAATCACCGTGAACGCCCAAGCATCTGCGCGCGCAC 37905
QY 315 TGAATTGCGGACGCGGAACCGGACCTAAGCCACCGGAGTGAATTCTCGTGCATGSCCAT 374
Db 37904 CAGGCTCGCAACGACACCGAGTGGCCGCGCACTTGAAGTTGATCTTAAGTCGACGCGCG 37845
QY 375 GTCCCTGCGGTACCCCGACATGCTTTGACGCGCGCTGGGCCATGCACTGCTGTTGA 434
Db 37844 CCAGCTGGCTTACCGGAGGGTTCTTTCAAGCGCGCATGGGCGATGCACTCGTCTGCA 37785
QY 435 GATGTCGGAACCGGACCGTGCATCCGGGAAATCTTTCAGTAATCTCAAAACCGGTGGCAT 494
Db 37784 GATCGTGAACAAGCCGCGCGCATTCGCGAGGTCCACCGAATCTTGAAACCGCGCGCGC 37725
QY 495 CCTCGGCGTCAACGAGTGTCAAAAGAAAGCGGCGGCGGATGCCGATGTCGCGGGA 554
Db 37724 GTTCGTCCTCGAGACATCATCTGGGTTGCACTCCGGAAGATACCGCGCGTTTG 37665
QY 555 CAGGTGGCCGACCGGCGCTTCGATCTGCTGCTGAGCAACTTCTGAAATCGCTGCGTGC 614
Db 37664 GACGGGCAAGACCGCCATACCTT-----GAACAGCTTCAAGCGCGTGGTCAAGCA 37614
QY 615 ACGGGGGTTTCAGATCTTCGATTGGAGAGACGTGTCTGAGGACCGCGTACTTCATGCC 674
Db 37613 AGCGGGGTTTCAGATTCTCGAAGTCAACGACTCACGSCACAGACCAAGCCAGGTGCAATGCTTC 37554
QY 675 GCAGTTCGCGGAAGAGCTGCTGCGGACAGCAAGGATCGCGGACAGGTACGCGCGCGGC 734
Db 37553 CTGGTACGTGAGAGATTGCTCCGAAACTCGATGAGCTGCGCGGCTGAGGCTTGGC 37494
QY 735 TGTGCGCGGCT 745
Db 37493 TGTGCGCACCT 37483

Search completed: August 29, 2005, 16:10:54
Job time : 3894 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 29, 2005, 12:38:52 ; Search time 585 Seconds

(without alignments)
8378.711 Million cell updates/sec

Title: US-10-069-353a-7

Perfect score: 828
Sequence: 1 ggtgtccaggtggtggtcacc.....cgcggaagccggtcgtcgtga 828

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	828	100.0	828	AAFB8318	AAFB8318 S. spinos
2	828	100.0	45624	AAFB8315	AAFB8315 S. spinos
3	828	100.0	50000	AAFB8312	AAFB8312 S. spinos
4	828	100.0	80161	AAZ21501	AAZ21501 DNA fragm
5	749.6	90.5	36538	ABV75558	ABV75558 Saccharop
6	170	20.5	846	AAZ54227	AAZ54227 Streptomy
7	170	20.5	52101	AAZ54217	AAZ54217 Streptomy
8	165.4	20.0	900	AAFB8324	AAFB8324 S. spinos
9	165.4	20.0	6085	AAFB8312	AAFB8312 S. spinos
10	165.4	20.0	45624	AAFB8315	AAFB8315 S. spinos
11	165.4	20.0	50000	AAFB8312	AAFB8312 S. spinos
12	165.4	20.0	80161	AAZ21501	AAZ21501 DNA fragm
13	159	19.2	36538	ABV75558	ABV75558 Saccharop
14	141.8	17.1	25681	ADJ53165	ADJ53165 Saccharop
15	133.2	16.1	852	AAFB8312	AAFB8312 S. spinos
16	133.2	16.1	852	AAFB8315	AAFB8315 S. spinos
17	133.2	16.1	53500	AAFB8312	AAFB8312 S. spinos
18	133.2	16.1	53500	AAFB8315	AAFB8315 S. spinos
19	130	15.7	1410	AAQ80524	AAQ80524 Oxioredu
20	122.8	14.8	2122	AAT70152	AAT70152 S. longisp

C 21	120.2	14.5	103599	4	ABX04971
C 22	112.2	13.6	12381	3	AAZ58381
C 23	93.2	11.3	1092	12	AD145320
C 24	89.4	10.8	1418	12	AD145598
C 25	84	10.1	1217	13	AD549291
C 26	83.2	10.0	828	3	AAZ5801
C 27	83.2	10.0	828	10	ADJ10220
C 28	74	8.9	942	10	ADZ62975
C 29	71.2	8.6	795	13	ADZ60017
C 30	71.2	8.6	110000	4	AA196882_41
C 31	71.2	8.6	110000	4	AA196883_41
C 32	67.6	8.2	67251	10	ADZ62995
C 33	67.2	8.1	2000	8	ADZ61938
C 34	64	7.7	2814	3	AAZ50960
C 35	63.2	7.6	1369	10	ABX12505
C 36	61	7.4	1131	11	ABD03718
C 37	61	7.4	1215	11	ABD03991
C 38	61	7.4	1215	11	ABD03909
C 39	60.6	7.3	7205	13	ADR84263
C 40	59	7.1	1862	4	AAH44261
C 41	59	7.1	2742	8	AAI61191
C 42	59	7.1	82746	8	AAI61324
C 43	58.4	7.1	2757	8	ACA38157
C 44	57.2	6.9	991	10	ABX12504
C 45	57.2	6.9	1152	10	ADK65628

ALIGNMENTS

RESULT 1
ID AAFB8318 standard; DNA; 828 BP.

AAFB8318,
28-AUG-2001 (first entry)

S. spinosa DNA fragment encoding ORF1, SEQ ID 7.

Forosamine; trimethylrhannose; polyketide synthase; biosynthesis;
KW spinosyn; polyketide aglycone; transgenic plant; insect resistance;
KM macrolide; insecticidal; O-methyltransferase; ds.

OS Saccharopolyspora spinosa.

PN DE19957268-A1.

PD 08-MAR-2001.

PF 29-NOV-1999; 99DE-01057268.

PR 27-AUG-1999; 99DE-01040596.

PA (FARB) BAYER AG.

Ederz G, Moehle V, Froede R, Velten R, Salas JA;

WPI; 2001-267102/28.

P-PSDB; AAB70948.

New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for recombinant production of insecticidal spinosyns and their derivatives.

Claim 7a; Page 102-104; 354pp; German.

This invention describes a novel method nucleic acid (i) and its encoded polypeptide (ii) containing at least one region that encodes an enzymatic activity involved in biosynthesis of spinosyns. (i) are used (i) to identify, inactivate or modulate genes involved in the biosynthesis of (ii); (ii) to generate a library of polyketide synthases; (iii) for adding forosamine or trimethylrhannose to a spinosyn or polyketide aglycone; and (iv) for recombinant production of the corresponding

ABX04971 S. cinam
AAZ58381 Streptomy
AD145320 Rice isop
AD145598 Wheat iso
AD549291 Bacterial
AAZ5801 S. lavend
ADJ10220 S. lavend
ADZ62975 Sorangium
ADZ60017 Bacterial
Continuation (42 o
Continuation (42 o
ADZ62995 Sorangium
ADZ61938 Rice gene
AAZ50960 A. haloph
ABX12505 CDNA enco
ABD03718 Pseudomon
ABD03991 Pseudomon
ABD03909 Pseudomon
ADZ84263 Aspergill
AAH44261 Physcomit
AAI61191 Actinosyn
AAI61324 Actinosyn
ACA38157 Prokaryot
ADK65628 CDNA enco
ADK65628 S. cerevis

enzymes, which are used for production of (II), their precursors or derivatives, including production of transgenic plants that express (II) and thus have increased resistance to insects. (I) are also useful as markers for sequencing of the Saccharopolyspora spinosa genome. (II) are macrolides with insecticidal, but not antibacterial, activity, and can also be used to raise specific antibodies, useful for identifying expression clones in a gene bank. Cells transformed with (I) may produce (II) at significantly increased levels or produce new derivatives of (II). This sequence, ORF 1, encodes an S. spinosa O-methyltransferase protein

Sequence 828 BP; 140 A; 262 C; 285 G; 141 T; 0 U; 0 Other;

Query Match 100.0%; Score 828; DB 4; Length 828;

Best Local Similarity 100.0%; Pred. No. 6.2e-168;

Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GTGTTGCCAGGTGGCGCACCAATCGCGAGGTTGGGCAATGATATCACTGTCACG 60
DB 1 GTGTTGCCAGGTGGCGCACCAATCGCGAGGTTGGGCAATGATATCACTGTCACG 60
QY 61 CGTTGCTGAACCTCGGTGGCGGGCCCTGCGGCATCAACGCGCTACCTGGGAGAC 120
DB 61 CGTTGCTGAACCTCGGTGGCGGGCCCTGCGGCATCAACGCGCTACCTGGGAGAC 120
QY 121 GACGGGCGGGCTTCTTCTGGCAGAGGCGCGCACCGGCTACCGACTTGTCCCGAACGG 180
DB 121 GACGGGCGGGCTTCTTCTGGCAGAGGCGCGCACCGGCTACCGACTTGTCCCGAACGG 180
QY 181 ACCGTGCTGATGCGGCGGCTGATGCTGATGAGGGGTGGGTTACCGGACCAACGAGC 240
DB 181 ACCGTGCTGATGCGGCGGCTGATGCTGATGAGGGGTGGGTTACCGGACCAACGAGC 240
QY 241 CTGGCGGTGCGCGCGACCAACCGGATCCAGATCACCGGATCAACGCTGACCGAGGTCGA 300
DB 241 CTGGCGGTGCGCGCGACCAACCGGATCCAGATCACCGGATCAACGCTGACCGAGGTCGA 300
QY 301 GTGGCCATGCGCGCTGATTCGCGACGCGAACCGGACTAAGCCACCGGGTGGACTTCTCG 360
DB 301 GTGGCCATGCGCGCTGATTCGCGACGCGAACCGGACTAAGCCACCGGGTGGACTTCTCG 360
QY 361 TGGCGTGAATGCTGCTGCTGCGGTAACCGGACATGCTTTGAGAGCGCGCTGGGCAATG 420
DB 361 TGGCGTGAATGCTGCTGCTGCGGTAACCGGACATGCTTTGAGAGCGCGCTGGGCAATG 420
QY 421 CAGTGGCTGTTGAGATGTCGGAACCGGACCGGTGCATCCGGGAAATCCTTGAGATATC 480
DB 421 CAGTGGCTGTTGAGATGTCGGAACCGGACCGGTGCATCCGGGAAATCCTTGAGATATC 480
QY 481 AAACCCGGTGGCATCTCGGCGTCAACCGAGTCTCAAAAGAAAGCGGCGCGGAGT 540
DB 481 AAACCCGGTGGCATCTCGGCGTCAACCGAGTCTCAAAAGAAAGCGGCGCGGAGT 540
QY 541 CCGGTGTCGGGGGACAGGTGGCGGACCGGCTTCCGATTCGCTGCTGAGCAACTTCTG 600
DB 541 CCGGTGTCGGGGGACAGGTGGCGGACCGGCTTCCGATTCGCTGCTGAGCAACTTCTG 600
QY 601 GAATGGCTGCTGACGCGGGGTTGAGATTCCTGATTTGGGAGAGAGTGCCTGCGAGACC 660
DB 601 GAATGGCTGCTGACGCGGGGTTGAGATTCCTGATTTGGGAGAGAGTGCCTGCGAGACC 660
QY 661 CGGTACTTCATGCGCGAAGTTCGCGAAGAGTCTGCTGGGACCAACGCGGATTCGCGAGC 720
DB 661 CGGTACTTCATGCGCGAAGTTCGCGAAGAGTCTGCTGGGACCAACGCGGATTCGCGAGC 720
QY 721 AGGTACGGGCGCGGCTGTGCGCGGCTGGGCGCGCGGCTGCTGAGATTAAGAAATATGCC 780
DB 721 AGGTACGGGCGCGGCTGTGCGCGGCTGGGCGCGCGGCTGCTGAGATTAAGAAATATGCC 780
QY 781 CACGACATGAGGCTATGCGATCTGAACGCGCGGAAAGCGGCTGCGACTGA 828
DB 781 CACGACATGAGGCTATGCGATCTGAACGCGCGGAAAGCGGCTGCGACTGA 828
```

RESULT 2

AAAF8315/c

ID AAAF8315 standard; DNA; 45624 BP.

AC AAAF8315;

DT 28-AUG-2001 (first entry)

DE S. spinosa DNA fragment SEQ ID 4.

KW Forosamine; trimethylrhinamose; polyketide synthase; biosynthesis;

KW spinosyn; polyketide aglycone; transgenic plant; insect resistance;

OS Saccharopolyspora spinosa.

PN DE19957268-A1.

PD 08-MAR-2001.

PF 29-NOV-1999; 99DE-01057268.

PR 27-AUG-1999; 99DE-01040596.

PA (FARB) BAYER AG.

DR Ebertz G, Moehrl V, Froede R, Velten R, Salas JA;

XX WPI; 2001-267102/28.

PT New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for recombinant production of insecticidal spinosyns and their derivatives.

PS Claim 7; Page 58-74; 354pp; German.

XX This invention describes a novel method nucleic acid (I) and its encoded CC polypeptide (II) containing at least one region that encodes an enzymatic CC activity involved in biosynthesis of spinosyns. (I) are used (i) to CC identify, inactivate or modulate genes involved in the biosynthesis of CC (II); (ii) to generate a library of polyketide synthases; (iii) for CC adding forosamine or trimethylrhinamose to a spinosyn or polyketide CC aglycone; and (iv) for recombinant production of the corresponding CC derivatives, which are used for production of (II), their precursors or CC derivatives, including production of transgenic plants that express (II) CC and thus have increased resistance to insects. (I) are also useful as CC markers for sequencing of the Saccharopolyspora spinosa genome. (II) are CC macrolides with insecticidal, but not antibacterial, activity, and can CC also be used to raise specific antibodies, useful for identifying CC expression clones in a gene bank. Cells transformed with (I) may produce CC (II) at significantly increased levels or produce new derivatives of CC (II). This sequence represents a genomic DNA fragment of the S. spinosa CC genome which contains the coding regions for proteins involved in CC forosamine and trimethylrhinamose biosynthesis

XX Sequence 45624 BP; 7933 A; 14913 C; 14940 G; 7838 T; 0 U; 0 Other;

Query Match 100.0%; Score 828; DB 4; Length 45624;

Best Local Similarity 100.0%; Pred. No. 7.8e-168;

Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GTGTTGCCAGGTGGCGCACCAATCGCGAGGTTGGGCAATGATATCACTGTCACG 60
DB 828 GTGTTGCCAGGTGGCGCACCAATCGCGAGGTTGGGCAATGATATCACTGTCACG 769
QY 61 CGTTGCTGAACCTCGGTGGCGGGCCCTGCGGCATCAACGCGCTACCTGGGAGAC 120
DB 768 CGTTGCTGAACCTCGGTGGCGGGCCCTGCGGCATCAACGCGCTACCTGGGAGAC 709
QY 121 GACGGGCGGGCTTCTTCTGGCAGAGGCGCGCACCGGCTACCGACTTGTCCCGAACGG 180
DB 708 GACGGGCGGGCTTCTTCTGGCAGAGGCGCGCACCGGCTACCGACTTGTCCCGAACGG 649
```

QY 181 ACCGTCCTGATGCGCGGCTTGCATGCTGATGTGGGGTGGGTAACGGAACACGCG 240
 XX |||||
 DB 648 ACCGTCCTGATGCGCGGCTTGCATGCTGATGTGGGGTGGGTAACGGAACACGCG 589
 QY 241 CTGCGCGTGGCGGCGACACGCGATCCAGATCACCGGATCACTGACGCGAGTGCAC 300
 XX |||||
 DB 588 CTGCGCGTGGCGGCGACACGCGATCCAGATCACCGGATCACTGACGCGAGTGCAC 529
 QY 301 GTGGCCATGCGCGTGGATTTGCGACGCGAACCGGACTTAAGCCACCGGGTGAATTCTTG 360
 XX |||||
 DB 528 GTGGCCATGCGCGTGGATTTGCGACGCGAACCGGACTTAAGCCACCGGGTGAATTCTTG 469
 QY 361 TGGCGTGAATGCCATGTCTTCCCTGCGTACCGGCAAACTTTGAGCGCGCTGGGCGATG 420
 XX |||||
 DB 468 TGGCGTGAATGCCATGTCTTCCCTGCGTACCGGCAAACTTTGAGCGCGCTGGGCGATG 409
 QY 421 CAGTCGCTGTTGAGATGTCCGAACCGGACCGGATCCGCGGAAATCTTGCAGTACTC 480
 DB 408 CAGTCGCTGTTGAGATGTCCGAACCGGACCGGATCCGCGGAAATCTTGCAGTACTC 349
 QY 481 AAACCCGTTGGCATCTCTGCGCGCTGACCGAGTGTCAAAAGAGACCGGCGCGGAGT 540
 DB 348 AAACCCGTTGGCATCTCTGCGCGCTGACCGAGTGTCAAAAGAGACCGGCGCGGAGT 289
 QY 541 CCGGTGTCCGGGAGACAGTGGCCGACCGGCTTGGATCTGCTGGCTGAGCAACTTCTG 600
 DB 288 CCGGTGTCCGGGAGACAGTGGCCGACCGGCTTGGATCTGCTGGCTGAGCAACTTCTG 229
 QY 601 GAATCGCTGCTGACGCGGAGTTCAGAGATCTTGCATTTGGAGAGACCTGTCTGAGAGAC 660
 DB 228 GAATCGCTGCTGACGCGGAGTTCAGAGATCTTGCATTTGGAGAGACCTGTCTGAGAGAC 169
 QY 661 CCGTACTTCAATGCGGAGTTCGCGAAGAGTCTGTCGCGACCAAGACCGGATCGCGAC 720
 DB 168 CCGTACTTCAATGCGGAGTTCGCGAAGAGTCTGTCGCGACCAAGACCGGATCGCGAC 109
 QY 721 AGGTACGGGCGCGGCTGTCCCGGCTGGCGCGCGCGCTGCGATTAAGAAATATGCG 780
 DB 108 AGGTACGGGCGCGGCTGTCCCGGCTGGCGCGCGCGCTGCGATTAAGAAATATGCG 49
 QY 781 CACGACATGGGCTATGAGATTTGACGCGCGCGGACCGGTCGGCTGA 828
 DB 48 CACGACATGGGCTATGAGATTTGACGCGCGCGGACCGGTCGGCTGA 1
 RESULT 3
 AAF8312
 ID AAF8312 standard; DNA: 50000 BP.
 XX
 AC AAF8312;
 XX
 DT 28-AUG-2001 (first entry)
 XX
 DE S. spinosa DNA fragment SEQ ID 1.
 XX
 KW Forosamine; trimethylthamose; polyketide synthase; biosynthesis;
 KM spinosyn; polyketide aglycone; transgenic plant; insect resistance;
 XX macrolide; insecticidal; ds.
 OS Saccharopolyspora spinosa.
 XX
 PN DE19957268-A1.
 XX
 PD 08-MAR-2001.
 PF 29-NOV-1999; 99DE-01057268.
 XX
 PR 27-AUG-1999; 99DE-01040596.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Eberz G, Moehle V, Froede R, Velten R, Salas JA,
 XX

DR WPI: 2001-267102/28.
 XX
 PT New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for
 recombinant production of insecticidal spinosyns and their derivatives.
 XX
 PS Claim 7; Page 14-31; 354pp; German.
 XX
 CC This invention describes a novel method nucleic acid (I) and its encoded
 CC polypeptide (II) containing at least one region that encodes an enzymatic
 CC activity involved in biosynthesis of spinosyns. (I) are used (i) to
 CC identify, inactivate or modulate genes involved in the biosynthesis of
 CC (II); (ii) to generate a library of polyketide synthases; (iii) for
 CC adding forosamine or trimethylthamose to a spinosyn or polyketide
 CC aglycone; and (iv) for recombinant production of the corresponding
 CC enzymes, which are used for production of (II), their precursors or
 CC derivatives, including production of transgenic plants that express (II)
 CC and thus have increased resistance to insects. (I) are also useful as
 CC markers for sequencing of the Saccharopolyspora spinosa genome. (II) are
 CC macrolides with insecticidal, but not antibacterial, activity, and can
 CC also be used to raise specific antibodies, useful for identifying
 CC expression clones in a gene bank. Cells transformed with (I) may produce
 CC (II) at significantly increased levels or produce new derivatives of
 CC (II). This sequence represents a genomic DNA fragment of the S. spinosa
 CC genome which contains the coding regions for proteins involved in
 CC forosamine, trimethylthamose and polyketide synthase biosynthesis
 CC
 SQ Sequence 50000 BP; 8484 A; 16290 C; 16535 G; 8691 T; 0 U; 0 Other;
 Query Match 100.0%; Score 828; DB 4; Length 50000;
 Best Local Similarity 100.0%; Pred. No. 7.9e-169;
 Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTGTTCCAGGTGGCGGACCAACATCCGACGAGTTGGGAGATGATGACCTGTCAGG 60
 DB 44797 GTGTTCCAGGTGGCGGACCAACATCCGACGAGTTGGGAGATGATGACCTGTCAGG 44856
 QY 61 CCGTGTCTGAATCGGTGCGCGGCGGCGCTGCGCATCCACACGCGTACTTGGAGAAC 120
 DB 44857 CCGTGTCTGAATCGGTGCGCGGCGGCGCTGCGCATCCACACGCGTACTTGGAGAAC 44916
 QY 121 GACGGGCGGGCTTCTCTGGAGAGAGCGCGCGACCGGCTACCGACCTTGTCCGGAACG 180
 DB 44917 GACGGGCGGGCTTCTCTGGAGAGAGCGCGCGACCGGCTACCGACCTTGTCCGGAACG 44976
 QY 181 ACCGTCTGATGAGCGGCGGCTTGCATGCTGATGTGGGGTGGGTAACGGAACACGCG 240
 DB 44977 ACCGTCTGATGAGCGGCGGCTTGCATGCTGATGTGGGGTGGGTAACGGAACACGCG 45036
 QY 241 CTGCGCGTGGCGGCGACACGCGATCCAGATCACCGGATCACTGACGCGAGTGCAC 300
 DB 45037 CTGCGCGTGGCGGCGACACGCGATCCAGATCACCGGATCACTGACGCGAGTGCAC 45096
 QY 301 GTGGCCATGCGCGTGGATTTGCGACGCGAACCGGACTTAAGCCACCGGGTGAATTCTTG 360
 DB 45097 GTGGCCATGCGCGTGGATTTGCGACGCGAACCGGACTTAAGCCACCGGGTGAATTCTTG 45156
 QY 361 TGGCGTGAATGCCATGTCTTCCCTGCGTACCGGCAAACTTTGAGCGCGCTGGGCGATG 420
 DB 45157 TGGCGTGAATGCCATGTCTTCCCTGCGTACCGGCAAACTTTGAGCGCGCTGGGCGATG 45216
 QY 421 CAGTCGCTGTTGAGATGTCCGAACCGGACCGGATCCGCGGAAATCTTGCAGTACTC 480
 DB 45217 CAGTCGCTGTTGAGATGTCCGAACCGGACCGGATCCGCGGAAATCTTGCAGTACTC 45276
 QY 481 AAACCCGTTGGCATCTCTGCGCGCTGACCGAGTGTCAAAAGAGACCGGCGCGGAGT 540
 DB 45277 AAACCCGTTGGCATCTCTGCGCGCTGACCGAGTGTCAAAAGAGACCGGCGCGGAGT 45336
 QY 541 CCGGTGTCCGGGAGACAGTGGCCGACCGGCTTGGATCTGCTGGCTGAGCAACTTCTG 600
 DB 45337 CCGGTGTCCGGGAGACAGTGGCCGACCGGCTTGGATCTGCTGGCTGAGCAACTTCTG 45396
 QY 601 GAATCGCTGCTGACGCGGAGTTCAGAGATCTTGCATTTGGAGAGACCTGTCTGAGAGAC 660

Db	45397	GAATGCTGCGAGCGGAGTTCGAGATTCGATTCGAGAGACGATTCGAGAGAC	45456
Qy	661	CGGACTTCATGCCGCGAGTTGCCGAAAGCTCGCTCGACACGACGCGGATCGCGAC	720
Db	45457	CGGACTTCATGCCGCGAGTTGCCGAAAGCTCGCTCGACACGACGCGGATCGCGAC	45516
Qy	721	AGTAAACGGCCGGGCTGTCGCGGCTGGCCCGCGGCTTCGATTATGAAATATGCC	780
Db	45517	AGTAAACGGCCGGGCTGTCGCGGCTGGCCCGCGGCTTCGATTATGAAATATGCC	45576
Qy	781	CACGACATGGGCTATGCGATTCTGACGCGCGGAGCCGGTCGGCTGA	828
Db	45577	CACGACATGGGCTATGCGATTCTGACGCGCGGAGCCGGTCGGCTGA	45624
RESULT 4			
ID	AA221501	standard; DNA; 80161 BP.	
AC	AA221501;		
XX	01-DEC-1999	(first entry)	
DE	DNA fragment of Saccharopolyspora spinosa containing biosynthetic genes.		
XX	Spinosyn biosynthetic enzyme; open reading frame; ORF;		
KW	insecticidal microlides; arachnid; nematode; insect; polyketide;		
KW	polyketide synthase; PKS; extender module; initiator module;		
KW	acyl transferase domain; AT; acyl carrier protein; ACP;		
KW	beta-ketosynthase domain; KS; KR; dehydratase domain; DH;		
XX	enoyl reductase domain; ER; beta-ketoreductase; insecticide; ss.		
OS	Saccharopolyspora spinosa.		
XX			
Key	Location/Qualifiers		
FT	complement(1135..1971)		
FT	/*tag= a		
FT	/product= "ORFL16"		
FT	/note= "Protein involved in transcription control"		
FT	2024..2791		
FT	/*tag= b		
FT	/product= "ORFL15"		
FT	/note= "Keto acyl reductase"		
FT	complement(3416..4165)		
FT	/*tag= c		
FT	/product= "spus"		
FT	/note= "Spinosyn biosynthesis protein S"		
FT	complement(4168..5325)		
FT	/*tag= d		
FT	/product= "spnr"		
FT	/note= "Spinosyn biosynthesis protein R"		
FT	complement(5363..6751)		
FT	/*tag= e		
FT	/product= "spnq"		
FT	/note= "Spinosyn biosynthesis protein Q"		
FT	7083..8450		
FT	/*tag= f		
FT	/product= "spnp"		
FT	/note= "Spinosyn biosynthesis protein P"		
FT	8967..10427		
FT	/*tag= g		
FT	/product= "spno"		
FT	/note= "Spinosyn biosynthesis protein O"		
FT	10436..11434		
FT	/*tag= h		
FT	/product= "spnn"		
FT	/note= "Spinosyn biosynthesis protein N"		
FT	complement(11530..12492)		
FT	/*tag= i		
FT	/product= "spnm"		
FT	/note= "Spinosyn biosynthesis protein M"		
FT	complement(12696..13547)		
CDS			
FT	/*tag= j		
FT	/product= "spnl"		
FT	/note= "Spinosyn biosynthesis protein L"		
FT	complement(13592..14785)		
FT	/*tag= k		
FT	/product= "spnk"		
FT	/note= "Spinosyn biosynthesis protein K"		
FT	complement(14799..16418)		
FT	/*tag= l		
FT	/product= "spnj"		
FT	/note= "Spinosyn biosynthesis protein J"		
FT	16556..17743		
FT	/*tag= m		
FT	/product= "spni"		
FT	/note= "Spinosyn biosynthesis protein I"		
FT	complement(17749..18501)		
FT	/*tag= n		
FT	/product= "spnh"		
FT	/note= "Spinosyn biosynthesis protein H"		
FT	complement(18541..19713)		
FT	/*tag= o		
FT	/product= "spng"		
FT	/note= "Spinosyn biosynthesis protein G"		
FT	20168..20995		
FT	/*tag= p		
FT	/product= "spnf"		
FT	/note= "Spinosyn biosynthesis protein F"		
FT	21111..28898		
FT	/*tag= q		
FT	/product= "spna"		
FT	/note= "Spinosyn biosynthesis protein A"		
FT	/transl_except= (Pos:26940..26942, aa:Pro)		
FT	21126..22379		
FT	/*tag= r		
FT	/note= "Beta-ketosynthase domain (KSI): part of the		
FT	initiator module"		
FT	22692..23669		
FT	/*tag= s		
FT	/note= "Acyl transferase domain (ATI): part of the		
FT	initiator module"		
FT	23793..24041		
FT	/*tag= t		
FT	/note= "Acyl carrier protein domain (ACPI): part of the		
FT	initiator module"		
FT	24102..25349		
FT	/*tag= u		
FT	/note= "Beta-ketosynthase domain (KSI): part of extender		
FT	module 1"		
FT	25683..26684		
FT	/*tag= v		
FT	/note= "Acyl transferase domain (ATI): part of extender		
FT	module 1"		
FT	27582..28121		
FT	/*tag= w		
FT	/note= "Beta-ketoreductase domain (KRI): part of extender		
FT	module 1"		
FT	28404..28649		
FT	/*tag= x		
FT	/note= "Acyl carrier protein domain (ACPI): part of		
FT	extender module 1"		
FT	29024..30295		
FT	/*tag= z		
FT	/note= "Beta-ketosynthase domain (KS2): part of extender		
FT	module 2"		
FT	30629..31621		
FT	/*tag= aa		
FT	/note= "Acyl transferase domain (AT2): part of extender		
FT	module 2"		
FT	31697..32254		
FT	/*tag= ab		
FT	/note= "Dehydratase domain (DH2): part of extender module		
FT	2"		
FT	33035..34072		
CDS			
FT	misc feature		
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			


```
FT      /tag= ac
FT      /note= "hnoyl" reductase domain (BR2) : part of extender
FT      module 2"
FT      34082. .34621
FT      /tag= ad
FT      /note= "Beta-ketoreductase domain (KR2) : part of extender
FT      module 2"
FT      34886. .30295
FT      /tag= ae
FT      /note= "Acyl carrier protein domain (ACP2) : part of
FT      extender module 2"
FT      35419. .44931
FT      /tag= af
FT      /product= "spnc"
FT      /note= "Spinosyn biosynthesis protein C"
FT      35518. .36786
FT      /tag= ag
FT      /note= "Beta-ketosynthase domain (KS3) : part of extender
FT      module 3"
FT      37108. .38097
FT      /tag= ah
FT      /note= "Acyl transferase domain (AT3) : part of extender
FT      module 3"
FT      38916. .35374
FT      /tag= y
FT      /product= "spnc"
FT      /note= "Spinosyn biosynthesis protein C"
FT      38992. .39528
FT      /tag= ai
FT      /note= "Beta-ketoreductase domain (KR3) : part of extender
FT      module 3"
FT      39790. .40035
FT      /tag= aj
FT      /note= "Acyl carrier protein domain (ACP3) : part of
FT      extender module 3"
FT      40102. .43373
FT      /tag= ak
FT      /note= "Beta-ketosynthase domain (KS4) : part of extender
FT      module 4"
FT      41713. .42705
FT      /tag= al
FT      /note= "Acyl transferase domain (AT4) : part of extender
FT      module 4"
FT      43615. .44157
FT      /tag= am
FT      /note= "Beta-ketoreductase domain (KR4) : part of extender
FT      module 4"
FT      44431. .44676
FT      /tag= an
FT      /note= "Acyl carrier protein domain (ACP4) : part of
FT      extender module 4"
FT      44966. .59752
FT      /tag= ao
FT      /product= "spnd"
FT      /note= "Spinosyn biosynthesis protein D"
FT      45077. .46348
FT      /tag= ap
FT      /note= "Beta-ketosynthase domain (KS5) : part of extender
FT      module 5"
FT      46691. .47674
FT      /tag= aq
FT      /note= "Acyl transferase domain (AT5) : part of extender
FT      module 5"
FT      47753. .48310
FT      /tag= ar
FT      /note= "Dehydratase domain (DH5) : part of extender module
FT      5"
FT      49226. .49771
FT      /tag= as
FT      /note= "Beta-ketoreductase domain (KR5) : part of extender
FT      module 5"
FT      50009. .50254
FT      misc_feature
```

```
Query Match      100.0%; Score 828; DB 2; Length 80161;
Best Local Similarity 100.0%; Pred. No. 8,1e-168;
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GTGTTGCCAGTGTGGCGCAACATCGACGAGTTTGGCAGATGTATGACTGTGTACG 60
Db      20168 GTGTTGCCAGTGTGGCGCAACATCGACGAGTTTGGCAGATGTATGACTGTGTACG 20227

Qy      61 CCGTTGCTGAACCTCGGTGCGGGCGGGCCCTTCGCCATTCACACCGGCTATCGGAGAAC 120
Db      20228 CCGTTGCTGAACCTCGGTGCGGGCGGGCCCTTCGCCATTCACACCGGCTATCGGAGAAC 20287

Qy      121 GACGGGCGGGCTTCCCTGGCAGCAGGCGCGCCGACCGGCTCAACGACCTTGTCCCGAACCG 180
Db      20288 GACGGGCGGGCTTCCCTGGCAGCAGGCGCGCCGACCGGCTCAACGACCTTGTCCCGAACCG 20347

Qy      181 ACCGTGCTGATGAGCGCGCTTGACTGCTGATGTGTGGGTGCGGTAACGACACACGCG 240
Db      20348 ACCGTGCTGATGAGCGCGCTTGACTGCTGATGTGTGGGTGCGGTAACGACACACGCG 20407

Qy      241 CTGCGCGTTCGCGCGCGCAACCGCATCCAGATCAACCGCATCAACCGGTGACCTTCTCG 300
Db      20408 CTGCGCGTTCGCGCGCGCAACCGCATCCAGATCAACCGCATCAACCGGTGACCTTCTCG 20467

Qy      301 GTGGCCATGCGCGCTGATTGGCGGACGGAACGCGGACTAAGCCACCGGCTGACCTTCTCG 360
Db      20468 GTGGCCATGCGCGCTGATTGGCGGACGGAACGCGGACTAAGCCACCGGCTGACCTTCTCG 20527

Qy      361 TGGCTGATGCCATGTCCTTCGCGGTAACCGGCAATGCTTTGACGCGCGCTGGGCGCATG 420
Db      20528 TGGCTGATGCCATGTCCTTCGCGGTAACCGGCAATGCTTTGACGCGCGCTGGGCGCATG 20587

Qy      421 CAGTCCCTGTTGGAAGATGTCGCCAACCGGACCGTGCCTCCGGAAATCCTTGAGATATC 480
Db      20588 CAGTCCCTGTTGGAAGATGTCGCCAACCGGACCGTGCCTCCGGAAATCCTTGAGATATC 20647

Qy      481 AAACCCGGTGGGACATCTCGGCGGTACCGAGGTGTCTAAACGAAACGGGCGCGGGATG 540
Db      20648 AAACCCGGTGGGACATCTCGGCGGTACCGAGGTGTCTAAACGAAACGGGCGCGGGATG 20707

Qy      541 CCGGTGTCCGGGAGCAGGTGGCGGACCGGCGCTTTGGATCTGCTGCTGAGCAATCTTCTG 600
Db      20708 CCGGTGTCCGGGAGCAGGTGGCGGACCGGCGCTTTGGATCTGCTGCTGAGCAATCTTCTG 20767

Qy      601 GAATCCCTGCGTGCACGCGGGGTTGAGATCTTGATTTGGAGAGACCTGTCTGAGAGACC 660
Db      20768 GAATCCCTGCGTGCACGCGGGGTTGAGATCTTGATTTGGAGAGACCTGTCTGAGAGACC 20827

Qy      661 CGGTACTTCAATGCGGAGTTGCGGAAAGCTGCTGCGCACGACGACGGGATCGCGGAC 720
Db      20828 CGGTACTTCAATGCGGAGTTGCGGAAAGCTGCTGCGCACGACGACGGGATCGCGGAC 20887

Qy      721 AGGTACGGGCGGCTGTGCGCGGCTGGGCGCGCGGCTGTGCAATTATGAAATATGCC 780
Db      20888 AGGTACGGGCGGCTGTGCGCGGCTGGGCGCGCGGCTGTGCAATTATGAAATATGCC 20947

Qy      781 CACGACATGGGCTATGCGATTCTGACGCGCGCGGAACCGGTCGGCTGCA 828
Db      20948 CACGACATGGGCTATGCGATTCTGACGCGCGCGGAACCGGTCGGCTGCA 20995

RESULT 5
ABV75558/c
ID      ABV75558 standard; DNA: 36538 BP.
XX      AC      ABV75558;
XX      DT      22-JAN-2003 (first entry)
XX      DE      Saccharopolyspora butenyl-spinosyn biosynthetic gene cluster 2.
XX      KM      Butenyl; biosynthetic enzyme; PKS; polyketide synthetase; macrolide;
          metabolite; spinosyn; gene; ds.
```

```
XX OS Saccharopolyspora sp.
XX Key Location/Qualifiers
FH CDS complement(114. .938)
FT /tag= a
FT /product= "buseF"
FT /note= "No start codon given"
FT CDS 1389. .2561
FT /tag= b
FT /product= "buseG"
FT CDS 2601. .3353
FT /tag= c
FT /product= "buseH"
FT CDS complement(3359. .4546)
FT /tag= d
FT /product= "buseI"
FT CDS 4684. .6303
FT /tag= e
FT /product= "buseJ"
FT CDS 6317. .7510
FT /tag= f
FT /product= "buseK"
FT CDS 7553. .8406
FT /tag= g
FT /product= "buseL"
FT CDS /note= "No start codon given"
FT 8640. .9572
FT /tag= h
FT /product= "buseM"
FT CDS /note= "No start codon given"
FT complement(9668. .10666)
FT /tag= i
FT /product= "buseN"
FT CDS complement(10675. .12135)
FT /tag= j
FT /product= "buseO"
FT CDS complement(12864. .14177)
FT /tag= k
FT /product= "buseP"
FT CDS 14627. .15970
FT /tag= l
FT /product= "buseQ"
FT CDS /note= "No start codon given"
FT 16008. .17144
FT /tag= m
FT /product= "buseR"
FT CDS /note= "No start codon given"
FT 17168. .17917
FT /tag= n
FT /product= "buseS"
FT CDS complement(18520. .19932)
FT /tag= o
FT /product= "ORF LI"
FT CDS complement(19978. .20488)
FT /tag= p
FT /product= "ORF LII"
FT CDS /note= "No start codon given"
FT complement(20536. .21033)
FT /tag= q
FT /product= "ORF LIII"
FT CDS /note= "No start codon given"
FT 21179. .21925
FT /tag= r
FT /product= "ORF LIV"
FT CDS complement(22671. .23453)
FT /tag= s
FT /product= "ORF LVI"
FT CDS complement(23687. .24886)
FT /tag= t
FT /product= "ORF LVII"
FT CDS complement(26177. .26923)
FT /tag= u
```

```
FT /product= "ORF LVIII"
FT /note= "No start codon given"
FT CDS 27646. .28476
FT /tag= v
FT /product= "ORF LIX"
FT /note= "No start codon given"
FT WO200279477-A2.
FT 10-OCT-2002.
FT 28-MAR-2002; 2002WO-US009968.
FT 30-MAR-2001; 2001US-0280175P.
FT (DOWC ) DOW AGROSCIENCES LLC.
FT Hahn DR, Jackson UD, Bullard BS, Gustafson GD, Waldron C;
FT Mitchell JC;
FT WPI; 2003-058434/05.
FT P-P8DB; ABP57683, ABP57684, ABP57685, ABP57686, ABP57687, ABP57688,
FT ABP57689, ABP57690, ABP57691, ABP57692, ABP57693, ABP57694, ABP57695,
FT ABP57696, ABP57697, ABP57698, ABP57699, ABP57700, ABP57701, ABP57702,
FT ABP57703, ABP57704.
FT New butenyl-spinosyn biosynthetic genes, useful for increasing the
FT production of butenyl-spinosyn insecticidal macrocides, or for changing
FT the metabolites or products produced by spinosyn-producing
FT microorganisms.
FT Claim 2; Page 99-119; 218pp; English.
XX PS
XX CC The invention relates to a novel DNA molecule comprising a DNA sequence
XX CC that encodes a butenyl-spinosyn biosynthetic enzyme, a butenyl-spinosyn
XX CC PKS (polyketide synthetase) domain, or a spinosyn PKS module. The butenyl
XX CC -spinosyn biosynthetic genes are useful for increasing the production of
XX CC butenyl-spinosyn insecticidal macrocides. The genes are also useful for
XX CC changing the metabolites or products produced by spinosyn-producing
XX CC microorganisms. The present sequence represents a DNA molecule encoding
XX CC butenyl-spinosyn biosynthetic enzymes
XX SQ
XX Sequence 36538 BP; 6867 A; 12266 C; 11182 G; 6223 T; 0 U; 0 Other;
Query Match 90.5%; Score 749.6; DB 10; Length 36538;
Best Local Similarity 94.1%; Pred. No. 4,88-151;
Matches 779; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 GTGTTGCCAGGTGCGCAGCAACAATCGAGCAGTTGGGCAATGATGATGACCTGTCAAG 60
DB 938 GTGTTGCCAGGTGCGCAGTCCCAACATCGCAGCAGGTCCGGCGAGATGTATGACTGTGTCAC 879
QY 61 CCGTTGCTGAACTGCGTCCGCGGCGGCCCTGCGCCCAATCCACCGCGCTACTGGGAAC 120
DB 878 CCGTTGCTGAACTGCGTCCGCGGCGGCCCTGCGCCCAATCCACCGCGCTACTGGGAAC 819
QY 121 GACGCGCGGCGCTTCTTCGACGACGACGCGCGACCGCGCTCAACGACTTGTGCGCGAAACG 180
DB 818 GACGCGCGGCGATTCTCTGACGACGACGCGCGCGACCGCGCTCAACGACTTGTGCGCGAAACG 759
QY 181 ACCGTGCTCATGCGCGCGCTTTCGACTGCTGATGTGGGTGCGGTACCGACGACGACGCG 240
DB 758 ACCGTGCTCATGCGCGCGCTTTCGACTGCTGATGTGGGTGCGGTACCGACGACGACGCG 699
QY 241 CTGCGCGTCCGCGCGCGCAACGCGATCCAGTCAACGCGGCAATCAACGTCAGCAGGTGCA 300
DB 698 CTGCGCGTCCGCGCGCGCAACGCGATCCAGTCAACGCGGCAATCAACGTCAGCAGGTGCA 639
QY 301 GTGCGCAATCGCGCGCTGATGCGACGCGAAACGCGCACTAAAGCCACCGGCTGACTTCTCG 360
DB 638 GCTGCGCAATCGCGCTGATGCGACGCGCGAAACGCGCACTAAAGCCACCGGCTGACTTCTCG 579
QY 361 TGCCTCATGCGATGCTCCCTGCGGTACCGCGGACAAATGCTTTCAGCGCGCGCTGCGCATG 420
```


XX Streptomyces platensis.
OS
XX
FH Key Location/Qualifiers
FT CDS complement(67..3720)
FT /tag= a
FT /product= "ORF1 protein"
FT CDS 4092..5681
FT /tag= b
FT /product= "ORF2 protein"
FT /note= "No start codon"
FT /partial
FT CDS 5767..6018
FT /tag= c
FT /product= "ORF3 protein"
FT 6023..7993
FT /tag= d
FT /product= "ORF4 protein"
FT 8009..17587
FT /tag= e
FT /product= "ORF5 protein"
FT 17634..41714
FT /tag= f
FT /product= "ORF6 protein"
FT /note= "No start codon"
FT /partial
FT CDS 41772..47633
FT /tag= g
FT /product= "ORF7 protein"
FT /note= "No start codon"
FT /partial
FT CDS 47635..49890
FT /tag= h
FT /product= "ORF8 protein"
FT 49922..50938
FT /tag= i
FT /product= "ORF9 protein"
FT 51234..52079
FT /tag= j
FT /product= "ORF10 protein"
XX
XX W0200288176-A2.
XX
XX 07-NOV-2002.
XX
XX 26-APR-2002; 2002MO-CA000591.
XX
XX 26-APR-2001; 2001US-0286346P.
XX
XX (ECOP-) ECOPIA BIOSCIENCES INC.
XX
XX Farnet CM, Zazopoulos E, Staffa A, Yang X;
XX
XX MPI, 2003-201222/19.
XX P-PBDB; AAE35485, AAE35486, AAE35487, AAE35488, AAE35489,
XX AAE35490, AAE35491, AAE35492, AAE35493.
XX
XX Novel isolated or purified polypeptide involved in biosynthesis of
XX polyketide dorriginocin or polyketide lactimidomycin, useful for preparing
XX dorriginocin or lactimidomycin.
XX
XX Claim 1; Page 85-113; 312pp; English.
XX
XX The invention relates to novel proteins involved in the biosynthesis of
XX polyketide dorriginocin (DORR) or lactimidomycin (LACT) biosynthesis by
XX microorganisms. Sequences of the invention allow direct manipulation of
XX dorriginocin, lactimidomycin and related chemical structures via chemical
XX engineering of the enzymes involved in the biosynthesis of dorriginocin and
XX lactimidomycin. They are useful for introducing chemical modifications into
XX normally inert positions that permit subsequence chemical modifications
XX and facilitate the development of polyketides. The genes and proteins of
XX the invention can also be used to generate a focused library of analogues
XX around a polyketide lead candidate to fine-tune the compound for optimal

CC properties. They are useful for generating antibodies specific for the
CC polyketide biosynthesis. The present sequence is Streptomyces platensis
CC subspecies rosaceus DORR DNA
XX
SQ Sequence 52101 BP; 7364 A; 20113 C; 17894 G; 6730 T; 0 U; 0 Other;
Query Match 20.5%; Score 170; DB 10; Length 52101;
Best Local Similarity 59.4%; Pred. No. 6.3e-27;
Matches 328; Conservative 0; Mismatches 215; Indels 9; Gaps 2;
QY 5 TCCCAAGTGGCGGCACCAACATGACAGCGAGTTGGCGAGATGTATGACTGTGACGCCGT 64
DB 51268 TCCTCCGTCCTCCGCCCTCCCGAGAGTGGACACTTACGACCGCTCACCGCAC 51327
QY 65 TGCTGAATCGATCGATCGCGGCGGCCCTCGGCATTCACACGAGTACTGGAGAACGACG 124
DB 51328 TGGACACCGAAGCGCGCGCG-----GCACTTCACCTGTGGTACTGGAGACGTGACG 51381
QY 125 GCGCGGCTTCTCTGGC--AGCAGGCGCGCGACCGGCTCACCGACTTGTGCGCGAACGGA 181
DB 51382 ACAACGACACCCCGCTGTGGAAAGGCGCGACCGGCTCACCGACACGATGACCGACGCC 51441
QY 182 CCGTCTGATGGCGGCGCTTGCATGCTTCATGTGGGTGGGTATACCGGACACCAACGACGC 241
DB 51442 TCGGATGACACGAGGACGAGGGGTCTCTGACAGTGGGTGGAGTGGCGGACCGGCA 51501
QY 242 TCGCGTGGCGCGGACACGAGATCCAGATCACCGGATCACCGTCCAGCGGATGACG 301
DB 51502 TCGGATGCGCGCGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 51561
QY 302 TGGCCATGCGCGCTGATTTGGCGACGCGAACGCGGACTAAGCACCGGATGAGTTCTTCTGT 361
DB 51562 TCGCCCGCGCACCGCGCTCCGCGAGGGCGCGCGCGCTGAGGACCGCGTGAATTCCGCG 51621
QY 362 GCGTCGATGATGATGCTCCGCGCTACCGCGACCATGCTTTGAGACGCGCGCGCGCATGCG 421
DB 51622 ACGCGAGCGCATGAACTGCTCTTCCCGACGACTCTTTCGAGCGCGCGCATTCGCCATCG 51681
QY 422 AGTCGCTGTGAGATGTCGCAACCGGACCGGACGCTGCATTCGGGAATCTTTCGATGACTCA 481
DB 51682 AGTCGATTTCCACATGCGCGACCGCGGACCGGATCTCTCGCGAATCCGCGCGTACTGC 51741
QY 482 AACCGGTGCGATCTCGGCGCTACCGAGTGTCAACGAGAACGGGCGCGCGGATGC 541
DB 51742 GCCCGGCGCGCGCGCTGCTCTTCCGACGACTTCTTCGAGCGCGCGCTCCCGCGAGA 51801
QY 542 CGGTGTCCGGGG 553
DB 51802 AGCAGCCCGCG 51813
RESULT 8
AAFB8324
ID AAFB8324 standard; DNA; 900 BP.
XX
XX AAFB8324;
XX
XX 28-AUG-2001 (first entry)
XX
XX S. spinosa DNA fragment encoding ORF6, SEQ ID 19.
XX
XX
XX Forosamine; trimethylxanthose; polyketide synthase; biosynthesis;
XX spinosyn; polyketide aglycone; transgenic plant; insect resistance;
XX macrolide; insecticidal; O-methyltransferase; ds.
OS Saccharopolyspora spinosa.
XX
XX DE19957268-A1.
XX
XX 08-MAR-2001.
XX
XX 29-NOV-1999; 99DF-01057268.
XX

PR 27-AUG-1999; 99DE-01040596.
 XX (FARB) BAYER AG.
 PA
 XX
 PI Eberz G, Moehle V, Froede R, Velten R, Salas JA;
 XX WPI, 2001-267102/28.
 DR P-PSDB; AAB70954.
 XX
 PT New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for
 PT recombinant production of insecticidal spinosyns and their derivatives.
 XX
 PS Claim 7a; Page 126-127; 354pp; German.
 XX
 CC This invention describes a novel method nucleic acid (I) and its encoded
 CC polypeptide (II) containing at least one region that encodes an enzymatic
 CC activity involved in biosynthesis of spinosyns. (I) are used (i) to
 CC identify, inactivate or modulate genes involved in the biosynthesis of
 CC (II); (ii) to generate a library of polyketide synthases; (iii) for
 CC adding forosamine or trimethylamine to a spinosyn or polyketide
 CC aglycone; and (iv) for recombinant production of the corresponding
 CC enzymes, which are used for production of (II), their precursors or
 CC derivatives, including production of transgenic plants that express (II)
 CC and thus have increased resistance to insects. (I) are also useful as
 CC markers for sequencing of the Saccharopolyspora spinosa genome. (II) are
 CC macrolides with insecticidal, but not antibacterial, activity, and can
 CC also be used to raise specific antibodies, useful for identifying
 CC expression clones in a gene bank. Cells transformed with (I) may produce
 CC (II) at significantly increased levels or produce new derivatives of
 CC (II). This sequence, ORF 6, encodes an S. spinosa O-methyltransferase
 CC protein
 CC
 SQ Sequence 900 BP; 172 A; 292 C; 289 G; 147 T; 0 U; 0 Other;
 Query Match 20.0%; Score 165.4; DB 4; Length 900;
 Best Local Similarity 52.8%; Pred. No. 4.8e-26;
 Matches 386; Conservative 0; Mismatches 336; Indels 9; Gaps 1;
 QY 15 CGCACCACATCCGACAGGTTGGGAGATGTATGACCTGTGACGCGCTGTGAATC 74
 DB 21 CACACAGGCAATTCGACCGCGGATCAGTGAGATCATCTTCGATGCGGTGGCGACG 80
 QY 75 GGTGCGGGGCGGCGCTCCCTGCGCCATCCACACGCGTACTGGAGAGACGCGGGGCTTC 134
 DB 81 GCGTCCCTTCACACGCGTACTGCGGCGGCGGTATCGGAGAGATGCCGCTCCACACC 140
 QY 135 CTGGCAGCAGGCGCGGCGGCGGCTCACCGACTTGTGTGCGGAACGACGCTGCTCATG 194
 DB 141 GTGATGAGATGCTGCGACCACTGACCACTGTTATGACAGAGCCGCGCTCCGTC 200
 QY 195 CGGCGTTGCACTGCTGATGTGGGTGCGGTACCGGACCAACAGCGCTGCGCGCG 254
 DB 201 CGAGAGGCACTGTTGCACTGCGGTGCGGCAATGGGACCGCTAGTCCGTGGCGATG 260
 QY 255 CGACACAGCGATCCAGATCACCGGCACTACCGTTCAGCAGAGTCAAGTGGCCATCGCGC 314
 DB 261 CGGCAAGGCGGCTTCAGATCACCGGAAACACCGTGAAGCCCGCATCTTCGCGCGCAC 320
 QY 315 TGATTGGGCAAGGCAAGCGGCACTAAGCAGCGGCGTGAATCTTCGCGCTGATGCGCAT 374
 DB 321 CAGGCTGCGCAAGCAGACCGGATGCGCGGCGGCTTGAATTCATGATGACGGCGC 380
 QY 375 GTCCCTGCGCTACCGGACCAATGCTTTCAGACCGCGCTGGGCGCATGCGATGCGTTGA 434
 DB 381 CCAAGCTGCGCTACCGGACCGGTTTCTTTCAGGCGCGCATGGCGCATGCGATCGCGTGA 440
 QY 435 GATGTCCGAACCGGACCGGTCATCCGGAATCTTTCAGATCTCAAAACCGGTGGCAT 494
 DB 441 GATGTGAGACCAAGCGCGCGGATCCCGAGATCCACCGATCTTGAAACCGCGGCGCG 500
 QY 495 CCGCGGCGTACCGGAGGTGCTCAAAACGAAAGCGGCGCGGAGTGCCTGTCGCGGGA 554
 DB 501 GTTGTCTCTCGAGACATCATCACTCGGGTTTCGACTCCGGAAGATACCGCGCGTTTG 560

QY 555 CAGTGGCCGACCGGCGCTTCGATCTGCTGCTGAGCACTTGTGAATGCGCTGTC 614
 DB 561 GACGGGACAGCGCGCCCAATCTT-----GAAACGCTTACGCGCGCTGTACGCA 611
 QY 615 ACCGGGTTTCGATCTCTGATTTGGAGAGACGTGTCTGAGAGACCCGCTACTTATGCC 674
 DB 612 AGCCGGGTTTCGATCTCTGATTTGGAGAGACGTGTCTGAGAGACCCGCTACTTATGCC 671
 QY 675 GCAATTCCGCGAAGAGCTCTGCGCACGACGAGATCGCGACAGGTACGGGCGGCG 734
 DB 672 CTGTATCGTCGACGAGTTGCTCCGAAAATCGATAGTCTCGCGCGCTGACCTCGCGC 731
 QY 735 TGTCCCGCGCT 745
 DB 732 TGTCCGCACTT 742
 RESULT 9
 ID AAT70153 standard; DNA; 6085 BP.
 AC AAT70153;
 XX
 DT 17-OCT-2003 (revised)
 DT 23-UTL-1997 (first entry)
 XX
 DE S.longisporoflavus staurosporin synthesis gene cluster 6.5kb fragment.
 XX Staurosporin; gene cluster; indole-carbazole alkaloid; ICA; antibiotic;
 KM antiproliferative; platelet aggregation; fungus; yeast;
 KM Ca2+/phospholipid-dependent serine/threonine protein kinase; ds.
 XX
 OS Streptomyces longisporoflavus; (strain R19).
 XX
 FH Key Location/Qualifiers
 FT CDS 378..1655
 FT /tag= a
 FT /label= Gene 1
 FT /note= "Encodes a protein containing 425 amino acids"
 FT 1747..2553
 FT /tag= b
 FT /label= Gene 2
 FT /note= "Encodes a protein containing 268 amino acids
 FT significantly similar to known S-adenosyl methionine-
 FT dependent methyl transferase"
 FT 2593..4011
 FT /tag= c
 FT /label= Gene 3
 FT /note= "Encodes a protein containing 472 amino acids"
 FT 4013..4959
 FT /tag= d
 FT /label= Gene 4
 FT /note= "Encodes a protein containing 328 amino acids"
 FT 5071..6085
 FT /tag= e
 FT /label= Gene 5 (part)
 FT /note= "Partial sequence for gene 5, full length protein
 FT contains 366 amino acids and is significantly similar to
 FT amino transferase enzymes, such as the Dnr J protein"
 PN MO9708323-A1.
 XX
 XX 06-MAR-1997.
 XX
 PF 19-AUG-1996; 96MO-BP003643.
 XX
 PR 30-AUG-1995; 95BP-00810534.
 XX
 XX (CIBA) CIBA GEIGY AG.
 PA
 PI Schupp T, Engel N, Bietenhader J, Toupet C, Pospiech A;

DR WPI, 1997-179280/16.
XX Indole-carbazole alkaloid biosynthesis gene cluster - especially coding
PT for the antibiotic staurosporin from Streptomyces longisporolavus.
XX
XX
PS Claim 7, Page 37-41; 55pp; English.
XX
XX The present sequence represents the 6.5kb PvuII fragment of Streptomyces
CC longisporolavus R19 which is involved in the biosynthesis of indole-
CC carbazole alkaloids (ICA). The sequence contains five functional
CC fragments as indicated in the features table. The DNA or a hybrid vector
CC containing it can be used to prepare an ICA or derivatives and
CC precursors, either by allowing production in previously incapable
CC organisms or by improving yields. In particular, the antibiotic
CC staurosporin can be produced. Staurosporin is known to have inhibitory
CC activity against fungi, yeasts, and Ca2+/phospholipid-dependent
CC serine/threonine protein kinases (PKs). Staurosporin also has
CC antiproliferative activity and can inhibit platelet aggregation. The
CC present sequence can also inactivate ICA biosynthesis genes and can be
CC used in PCR amplification. An advantage of this is that productivity of
CC staurosporin-synthesizing Streptomyces is improved over natural strains
CC yielding only low concentrations. (Updated on 17-OCT-2003 to standardise
CC OS field)
XX
XX
SQ Sequence 6085 BP, 882 A; 2374 C; 1992 G; 833 T; 0 U; 4 Other;

Query Match 20.0%; Score 165.4; DB 2; Length 6085;
Best Local Similarity 60.6%; Pred. No. 5.4e-26;
Matches 271; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

OY 90 CTGCGCCATCCACACGCGCTACTGTGGAGAACGACGCGGCGCTTCTTGCGACAGAGCGCGC 149
DB 1833 CGGCAACATCCACGCGGGTACTGTGGAGAACGACCGGAGGTCGATTCGCGAGCGCAC 1892
OY 150 CGACCGGCTCACCGGACCTTTGCGCGGAACGACCGTGTTCATGGCGCGCTTTCATCTCT 209
DB 1893 CGACCGGCTCACCGGACCTTTGCGCGGAACGACCGTGTTCATGGCGCGCTTTCATCTCT 1952
OY 210 CGATGTGGGGTTCGCGTACCGGACCAACGCGCTGCGCGCGGACCAACGCGATCCA 269
DB 1953 GGAAGTGGGCTTCGCGCATTCGCGTCCGCGCTTCAGATTCGCGGACGCGACGATCCG 2012
OY 270 GATCAACCGGCATCACCGTCAAGCAGAGTCAAGTGGCCATTCGCGCTGATTCGCGACGCGA 329
DB 2013 CGTCAACCGGCATCACCGTCAAGCAGCAGCAGGTCACCGAGGCGCGGACGCGGCGATGGA 2072
OY 330 AGCGGACCTAAGCCACCGGCGTGAAGTCTTCGCGCTGCATGCGCATGCTCCGCTGACCC 389
DB 2073 GTCCGATCCCGGGGCGGCGTCTCTTCGCGCTGCGGACCGCATGGAAGCTTCCCTTGA 2132
OY 390 GGACATGCTTTTCGACGCGCGCTGGGCGCATGCGCTGTGGAGATGTCGGAACCGGA 449
DB 2133 GGAAGTCTCTTTTCGACGCGGCGCTTTCGCGATGAGTCTGTGTCATTCGCGGACCAAC 2192
OY 450 CCGTCCATTCGCGGAATCTTTCGAGTACTCAACCGGCTGCGATCTTCGCGGCTACCGA 509
DB 2193 ACCCGCGCTCAAGGATTCACCGGCGTGTCTCGCGCGCGCGCGCTTCATTCATCCGA 2252
OY 510 GGTGCTCAAGAGAGGCGGCGCGG 536
DB 2253 CCGTGTCAAGCAGACGCGGTTCACCGG 2279

RESULT 10
AAF88315 standard: DNA; 45624 BP.
XX AAF88315;
XX AC
XX 28-AUG-2001 (first entry)
XX
XX S. spinosa DNA fragment SEQ ID 4.
DB XX

KM Forosamine; trimethylrhinamose; polyketide synthase; biosynthesis;
XX spinosyn; polyketide aglycone; transgenic plant; insect resistance;
KM macrolide; insecticidal; da.
XX
XX
OS Saccharopolyspora spinosa.
XX
XX DE19957268-A1.
XX
XX
PD 08-MAR-2001.
XX
XX
PF 29-NOV-1999; 99DE-01057268.
XX
XX
PR 27-AUG-1999; 99DE-01040596.
XX
XX
PA (FARB) BAYER AG.
XX
XX
PI Ebertz G, Moehrle V, Froede R, Velten R, Salas JA;
XX
XX
DR WPI, 2001-267102/28.
XX
XX
XX

New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for recombinant production of insecticidal spinosyns and their derivatives.

Claim 7, Page 58-74; 354pp; German.

This invention describes a novel method nucleic acid (I) and its encoded polypeptide (II) containing at least one region that encodes an enzymatic activity involved in biosynthesis of spinosyn. (I) are used (i) to identify, inactivate or modulate genes involved in the biosynthesis of (II); (ii) to generate a library of polyketide synthases; (iii) for adding forosamine or trimethylrhinamose to a spinosyn or polyketide aglycone; and (iv) for recombinant production of the corresponding enzymes, which are used for production of (II), their precursors or derivatives, including production of transgenic plants that express (II) and thus have increased resistance to insects. (I) are also useful as markers for sequencing of the Saccharopolyspora spinosa genome. (II) are CC macrocyclic with insecticidal, but not antibacterial, activity, and can CC also be used to raise specific antibodies, useful for identifying CC expression clones in a gene bank. Cells transformed with (I) may produce CC (II) at significantly increased levels or produce new derivatives of CC (II). This sequence represents a genomic DNA fragment of the S. spinosa CC genome which contains the coding regions for proteins involved in CC forosamine and trimethylrhinamose biosynthesis
XX
XX
SQ Sequence 45624 BP, 7933 A; 14913 C; 14940 G; 7838 T; 0 U; 0 Other;

Query Match 20.0%; Score 165.4; DB 4; Length 45624;
Best Local Similarity 52.8%; Pred. No. 6.1e-26;
Matches 386; Conservative 0; Mismatches 336; Indels 9; Gaps 1;

OY 15 CGACCAACATCGACGAGGTGGGAGATGATGACCTGTGTCAGGCGCTTGTGTAATC 74
DB 7421 CACACAGGCAATTCGACCGCGATCAGTGTGATTCATCTTCATGCGGCGACCG 7480
OY 75 GGTGCGGGGCGGCGCTTCGCGCATTCACACGCTACTGTGAGAACGACGCGGCGCTTC 134
DB 7481 GCGTCCCTTCGACCAACGCTTACTGGGCGGCGGATTCGGAGAGATCCGCTGCCAACCC 7540
OY 135 CTGCGACAGGCGCGCGACCGGCTCACCGACCTTTGCGCGGAACGACCGGCTTCGATGG 194
DB 7481 GCGTCCCTTCGACCAACGCTTACTGGGCGGCGGATTCGGAGAGATCCGCTGCCAACCC 7540
OY 7541 GTGTCGAGATGTCGACCAACGCTGACCGACCTGTTATGACCAAGGCGCGGCTTCGCTTC 7600
DB 195 CCGCGTTGACTGCTCGATGTGGGGTTCGCGTACCGGACCAACGACGCTGCGCTGCGCG 254
DB 7601 CGAGCGGCACTGTTTCACCTCGGCTTCGCGCAATGGGCAACCGCTATCCGTGGCGCATG 7660
OY 255 CGACCAACGATCCAGTCAACCGGCTACCGGCTGACGCGAGGTGCAAGTGGCGCATCGCGC 314
DB 7661 CGCCAGCGGCTTGAATTCACCGGATCACCGGATCACCGGATCACCGGATCACCGGATCAC 7720
OY 315 TGATTTGGCAGCGAAGCGGACCTAAGCCACCGGCTGATCTTCGCGCGCTTCGATTCGCAT 374
DB 7721 CAGGCTCGCAACGACGCGGACTGCGGCGAGCTTGAAGTTGATCTAAGTCAGCGGCGC 7780


```

FT misc_feature 37108..38097
FT /tag= ah
FT /note= "Acyl transferase domain (AT3) : part of extender
FT module 3"
FT CDS 38916..35374
FT /tag= y
FT /product= "apnc"
FT /note= "Spinosyn biosynthesis protein C"
FT 38992..39528
FT /tag= a1
FT /note= "Beta-ketoreductase domain (KR3) : part of extender
FT module 3"
FT 39790..40035
FT /tag= a1
FT /note= "Acyl carrier protein domain (ACP3) : part of
FT extender module 3"
FT 40102..41373
FT /tag= ak
FT /note= "Beta-ketosynthase domain (KS4) : part of extender
FT module 4"
FT 41713..42705
FT /tag= a1
FT /note= "Acyl transferase domain (AT4) : part of extender
FT module 4"
FT 43615..44157
FT /tag= am
FT /note= "Beta-ketoreductase domain (KR4) : part of extender
FT module 4"
FT misc_feature 44431..44676
FT /tag= an
FT /note= "Acyl carrier protein domain (ACP4) : part of
FT extender module 4"
FT 44966..59752
FT /tag= ao
FT /product= "apnd"
FT /note= "Spinosyn biosynthesis protein D"
FT 45077..46348
FT /tag= ap
FT /note= "Beta-ketosynthase domain (KS5) : part of extender
FT module 5"
FT 46691..47674
FT /tag= aq
FT /note= "Acyl transferase domain (AT5) : part of extender
FT module 5"
FT 47753..48310
FT /tag= ar
FT /note= "Dehydratase domain (DH5) : part of extender module
FT 5"
FT 49226..49771
FT /tag= as
FT /note= "Beta-ketoreductase domain (KR5) : part of extender
FT module 5"
FT 50009..50254
FT misc_feature

Query Match 20.0%; Score 165.4; DB 2; Length 80161;
Best Local Similarity 52.8%; Pred. No. 6.3e-26;
Matches 386; Conservative 0; Mismatches 336; Indels 9; Gaps 1;

QY 15 CGGACCAACATCCAGAGGTTGGGAGATGATGACCTGATCAGCGCGTTGCTGAATC 74
DB 13575 GACACAGGCAATTCGACCGCGATCAGGTGAGTCCATCTTGATCCGTTGGCGCACG 13516
QY 75 GGTGCGGCGCGCCCTGCGCATCCACACGAGCTACGAGAGAAACGCGCGGCTTC 134
DB 13515 GCGTCCCTCGACACAGGTTACTGCGCGCGGATCGGAGATGCCGTCACAC 13456
QY 135 CTGGCAGCAGCGCGCGCTCACACCACTTGTGCGCGAAGCAGCCTGTCTCAATG 194
DB 13455 GTGGTCGATGCTGCCACCACTGACCACTGTTCATGACAAAGCGCGCTCCGTC 13396
QY 195 CGGAGTTTCACTGCTGATGTGGGATGCGGTACCGAGCAACAGCGCTGCGCGTGGCG 254
DB 13395 CGAGCGCACTGTTCACTGGGCTGCGCAATGGCAAGCCCTAGTCCGTGCGGATG 13336

```

```

QY 255 GCACACCGCATCCAGATCACCGGCATCAGCTCAGCCAGGTGCAATGGCCATGCGGC 314
DB 13335 CGCCAGCGCGCTTGCAGTCAACCGGAATCACCGTGAAGCCAGCATCTCGCGCGCCAC 13276
QY 315 TGATTGCGCAGCGAAGCGGATCAAGCCACCGGTTGACTTCTGTGCTGATGCAAT 374
DB 13275 CAGGCTCCCAACGAGACCGGACTGCGCGGAGTCTTGTAGTTCATGATGACAGCGC 13216
QY 375 GTCCCTGCGGATCCCGGACAAATGCTTTCAGAGCGCGCCCTGGGCGATGCTGTGGA 434
DB 13215 CCAGCTGCCATACCGGACGAGTCTTTCAGGCGCGATGGGATGCAAGTCCGTGTGA 13156
QY 435 GATGTCGACCGACCGGCGCATCCGGAATCTTGTGAGTCAACTCAACCGGTGGAT 494
DB 13155 GATGTCGACGAGCGCGCGCGATCCGAGATCCAGGAATCTGGAACCGGAGCGCG 13096
QY 495 CCTCGCGTCAACCGAGTCTGCAACGAAACGAGCGCGGAGATCCGCTGTCGAGGA 554
DB 13095 GTTCGTCTCGGAGACATCATCATCTCGGTTTCGATCTCCGGAAGATACGCGCGGTTTG 13036
QY 555 CAGTGGCCGACCGGCTTTCGATCTGCTGCTGCTGAGCACTTCTGGAATGCTGCTGTC 614
DB 13035 GACGCGCACGACCGGCTTACCTT-----GAACAGCTTACGCGCGCTGTAAGCA 12985
QY 615 AGCGGAGTTTCAGATCTCGATTGGAGAGACGTGTCGAGAGACCGGTAATTGATGCC 674
DB 12984 AGCCGAGTTTCAGATTCTCGAAGTACAGCACTTCAAGCAGACAGACAGATGATGCTTC 12925
QY 675 GCAGTTCGCGGAAGTCTGCTGCGACACGACGAGATGCGGACAGTACGAGCGCGC 734
DB 12924 CTGGATCGTCAAGATGCTGCTCGGAACATGATAGCTCGCGCGTCAAGCTGCGCG 12865
QY 735 TGTGCGCGGCT 745
DB 12864 TGTGCGCACT 12854

RESULT 13
ABV75558
ID ABV75558 standard; DNA; 36538 BP.
XX
AC ABV75558;
XX
DT 22-JAN-2003 (first entry)
XX
DE Saccharopolyspora butenyl-spinosyn biosynthetic gene cluster 2.
XX
KW Butenyl: biosynthetic enzyme; PKS; polyketide synthetase; macrolide;
KW metabolite; spinosyn; gene; ds.
XX
OS Saccharopolyspora sp.
XX
XX
Key Location/Qualifiers
CDS complement(114..938)
FT /tag= a
FT /product= "busf"
FT /note= "No start codon given"
FT 1389..2561
FT /tag= b
FT /product= "busg"
FT 2601..3353
FT /tag= c
FT /product= "busH"
FT /product= "busH"
FT /tag= d
FT /product= "busJ"
FT 4684..6303
FT /tag= e
FT /product= "busJ"
FT 6317..7510
FT /tag= f
FT /product= "busK"

```

FT	CDS	7555. .8406
FT		/tag= g
FT		/product= "buaL"
FT	CDS	/note= "No start codon given"
FT		8640. .9572
FT		/tag= h
FT		/product= "buaM"
FT	CDS	/note= "No start codon given"
FT		complement (9668. .10666)
FT		/tag= i
FT		/product= "buaN"
FT	CDS	complement (10675. .12135)
FT		/tag= j
FT		/product= "buaO"
FT	CDS	complement (12864. .14177)
FT		/tag= k
FT		/product= "buaP"
FT	CDS	14627. .15970
FT		/tag= l
FT		/product= "buaQ"
FT		/note= "No start codon given"
FT	CDS	16008. .17144
FT		/tag= m
FT		/product= "buaR"
FT		/note= "No start codon given"
FT	CDS	17168. .17917
FT		/tag= n
FT		/product= "buaS"
FT	CDS	complement (18520. .19932)
FT		/tag= o
FT		/product= "ORF LI"
FT	CDS	complement (19978. .20488)
FT		/tag= p
FT		/product= "ORF LII"
FT		/note= "No start codon given"
FT	CDS	complement (20536. .21033)
FT		/tag= q
FT		/product= "ORF LIII"
FT		/note= "No start codon given"
FT	CDS	21179. .21925
FT		/tag= r
FT		/product= "ORF LIV"
FT	CDS	complement (22671. .23453)
FT		/tag= b
FT		/product= "ORF LVI"
FT	CDS	complement (23687. .24886)
FT		/tag= t
FT		/product= "ORF LVII"
FT	CDS	complement (26177. .26923)
FT		/tag= u
FT		/product= "ORF LVIII"
FT		/note= "No start codon given"
FT	CDS	27646. .28476
FT		/tag= v
FT		/product= "ORF LIX"
FT		/note= "No start codon given"
XX	W0200279477-A2.	
XX	EN	
XX	PD	10-OCT-2002.
XX	PF	28-MAR-2002; 2002WO-US009968.
XX	PR	30-MAR-2001; 2001US-0280175P.
XX	PA	(DOWC) DOW AGROSCIENCES LLC.
XX	PI	Hahn DR, Jackson JD, Bullard BS, Gustafson GD, Waldron C;
XX	PI	Mitchell JC;
XX	WP1	2003-058434/05.
DR	P-PSDB	ABP57683, ABP57684, ABP57685, ABP57686, ABP57687, ABP57688,
DR	ABP57689, ABP57690, ABP57691, ABP57692, ABP57693, ABP57694, ABP57695,	

Query Match	19.2%	Score 159;	DB 10;	Length 36538;
Best Local Similarity	52.3%	Pred. No. 1.4e-24;		
Matches 382;	Conservative 0;	Mismatches 340;	Indels 9;	Gaps 1;
Db	15	CGCACCAATCTCGAGAGCAGGATTGGGACAGATATATAGACTGTGTCAACGCGCTTGTGAATCTC	74	
Qy	7527	CGCACAGGCTTTTCCAAAGCGGGATCAAGTGAATCTTCAACGCGTTGGGGCAAG	7586	
Db	75	GATGCGGGGGCCCTTCGCGCATCCACACGCTACTGGGAGAACGACGGGGCGCTTC	134	
Qy	7587	GCGTGCCCTTCGACGACGAGTACTGAGCGGGGATCGGAGAGATCGGGGGCACAC	7646	
Db	135	CTGGAGAGGCGCGCGGACCGGCTCACCGACTTGTGCGCGAAGCGACCGGTCTGATGG	194	
Qy	7647	TTGTGTCGACGCTGTGCGACACTGACCGACTTGTTCATGACAAAGCGCGCTTCGCCC	7706	
Db	195	CGCGTTGACTGCTGATGTGTGGGTGCGGTACCGGACAAACAGCGCTGCGGTGCGCG	254	
Qy	7707	CGGAGCGCACTGTTTCGACTGTGGCTGTGGCAATGGGACGCCGTATGTCGCGGGACG	7766	
Db	255	CGACAGCGCATTCGAGATCACCGGATCACCGTACGACGAGTGGAAATGGCGATCGCGCG	314	
Qy	7767	CACCAAGCGCTTGAGTACACCGAATACCTGTAAAGCCGAAATCTCGCGCGCTAC	7826	
Db	315	TGATTTCGACAGCGCAACGCGGACTTAAGCCACCGGGTGAATCTTCGTGCTGATGTCAT	374	
Qy	7827	CAGGTGCGCAACGAGACCGGACTGTGGCGGACAGTCTTCGTTTCATGTATGTCAGCGCGC	7886	
Db	375	GTCCTGCGGATCCCGGACAAATGCTTTTCGACGCGCGCTGGGCGATGCAATGCTGTTTGA	434	
Qy	7887	CCGGTGTCCCTTACCCCGGAAGTTCTTTCACGCGGACATGGCGGATGAGTCCGTGTGATCA	7946	
Db	435	GATGTCGCAACCGGACCGGTGCATTCGCGGAAATCTTTCGAGTACTTCAAAACCGGTGGCAT	494	
Qy	7947	GATGTGTACCAAGCTGTGCGGATTCGCGGAGTTCACACGAAATCTTGGAAACCGGCGGCA	8006	
Db	495	CCTCGGCTCACCGAGGTGTCAAAAGAAAGCGGCGCGCGGAGTGCCTGTGTCGGGGA	554	
Qy	8007	GTTGTGTCTCGGGGACATCATCTCGTGTCTGCACTCCCGGAAGATGACGGGCGGTTTG	8066	
Db	555	CAGGTGCGGACCGGCGCTTTCGGAATCTGCTGCGGTGAGACCACTTGTGGAATGCTGCGTGC	614	
Qy	8067	GACCGGACGACCGGCGCATACCTT-----GAAAGCTTCAACCGCGCTGTGAAGGA	8117	
Db	615	AGCGGGGTTGAGATCTCGATTGGGAGAGAGTGTCTGAGAGACCGGTAATCTTCAATGCC	674	
Qy	8118	AGCGGGGTTGAGATCTTCTCAAGTCAACGACTCAACGCGCGGACGAGTGCATGTGTCTC	8177	
Db	675	GCAAGTTCGCGAAGAGCTCGCTGCGCACACGACGGAATGCGCGGACAGGTACGGGCGGCG	734	
Qy	8178	CTGGATGTGTCAGAGATTGCTCCGGGAAATCTGATGAGACTGCGCGGCTGTGAGCTCTGCGCG	8237	

QY 735 TGTCCGGGCT 745
 DB 8238 TGTCCGACCT 8248

RESULT 14
 ADJ53165
 ADJ53165 standard; DNA; 25681 BP.

ADJ53165;
 06-MAY-2004 (first entry)

Saccharothrix aerocolonigenes rebeccamycin biosynthetic gene cluster.
 Rebeccamycin biosynthetic gene cluster; rebeccamycin production;
 indolocarbazole production; ATCC39243; Streptomyces albus;
 antitumor agent; cytostatic; Gram positive bacterium; antibacterial;
 antibiotic; gene; ds.

Lechevalieria aerocolonigenes; ATCC39243.

Location/Qualifiers
 2..136
 /*tag= a
 /partial
 /product= "OrfD13 protein (SEQ ID NO:2)"
 /note= "No start codon given"
 complement(302..3313)
 /*tag= b
 /product= "OrfR5 protein (SEQ ID NO:3)"
 /function= "Regulatory protein"
 /transl_except= (pos:3310..3313,aa:Met)
 complement(3395..4027)
 /*tag= c
 /product= "OrfR4 protein (SEQ ID NO:4)"
 /function= "Dipeptidase"
 4402..5718
 /*tag= d
 /product= "OrfD1 protein (SEQ ID NO:5)"
 /function= "Sesterase"
 complement(5946..6347)
 /*tag= e
 /product= "OrfR3 protein (SEQ ID NO:6)"
 6581..7768
 /*tag= f
 /product= "OrfD2 protein (SEQ ID NO:7)"
 complement(7841..9106)
 /*tag= g
 /product= "OrfR2/ngt protein (SEQ ID NO:8)"
 /function= "N-glucosyltransferase"
 9316..10737
 /*tag= h
 /product= "OrfD3 protein (SEQ ID NO:9)"
 /function= "L-tryptophan oxidase"
 10734..13775
 /*tag= i
 /product= "OrfD4 protein (SEQ ID NO:10)"
 13772..15361
 /*tag= j
 /product= "OrfD5 protein (SEQ ID NO:11)"
 /function= "Monooxygenase"
 15358..16551
 /*tag= k
 /product= "OrfD6 protein (SEQ ID NO:12)"
 /function= "Cytochrome P450"
 /transl_except= (pos:15358..15360,aa:Met)
 16578..17399
 /*tag= l
 /product= "OrfD7 protein (SEQ ID NO:13)"
 /function= "Methyltransferase"
 /transl_except= (pos:16578..16580,aa:Met)
 17730..20501

FT /*tag= m
 FT /product= "OrfD8 protein (SEQ ID NO:14)"
 FT /function= "Regulatory protein"
 FT 20498..21010
 FT /*tag= n
 FT /product= "OrfD9 protein (SEQ ID NO:15)"
 FT /function= "Flavin reductase"
 FT 21007..22287
 FT /*tag= o
 FT /product= "OrfD10 protein (SEQ ID NO:16)"
 FT /function= "Membrane transport protein"
 FT 22271..23863
 FT /*tag= p
 FT /product= "OrfD11 protein (SEQ ID NO:17)"
 FT /function= "Tryptophan halogenase"
 FT complement(23933..25354)
 FT /*tag= q
 FT /product= "OrfR1 protein (SEQ ID NO:18)"
 FT /function= "Membrane transport protein"
 FT 25439..25681
 FT /*tag= r
 FT /partial
 FT /product= "OrfD12 protein (SEQ ID NO:19)"
 FT /function= "Regulatory protein"
 FT /note= "No stop codon given"
 FT /transl_except= (pos:25439..25441,aa:Met)

MO2003033706-A1.
 24-APR-2003.
 17-OCT-2002; 2002MO-ES000492.
 19-OCT-2001; 2001ES-00002312.
 (UYOV-) UNITV OVIEDO.
 Sanchez Reillo C, Fernandez Brana A, Salas Fernandez JA;
 Mendez Fernandez C;
 WPI; 2003-393533/37.
 P-PSDB; ADJ53166, ADJ53167, ADJ53168, ADJ53169, ADJ53170, ADJ53171,
 DR ADJ53172, ADJ53173, ADJ53174, ADJ53175, ADJ53176, ADJ53177, ADJ53178,
 DR ADJ53179, ADJ53180, ADJ53181, ADJ53182, ADJ53183.
 Production of indolo-carbazole antitumor agents, especially rebeccamycin
 or derivatives, by culturing host cells transformed with vector
 containing DNA from Saccharothrix aerocolonigenes.

Claim 2; SEQ ID NO 1; 76pp; Spanish.

The invention relates to a method for the production of indolocarbazoles
 in Streptomyces (especially Streptomyces albus) using rebeccamycin
 CC bioisynthetic genes isolated from Saccharothrix aerocolonigenes ATCC39243.
 CC The invention also relates to the Saccharothrix aerocolonigenes ATCC39243
 CC rebeccamycin biosynthetic gene cluster (ADJ53165) and sequences at least
 CC 80% homologous to it; nucleic acids encoding one or more rebeccamycin
 CC bioisynthetic proteins; the 18 rebeccamycin biosynthetic proteins
 CC (ADJ53166-ADJ53183) encoded by the gene cluster; and vectors and host
 CC cells comprising nucleic acid sequences of the invention.
 CC Indolocarbazoles (especially rebeccamycin) and their derivatives and
 CC precursors are useful as antitumor agents. Rebeccamycin also shows
 CC antibacterial activity against Gram positive bacteria such as
 CC Staphylococcus aureus, Micrococcus luteus and Streptococcus faecalis. The
 CC recombinant Streptomyces albus of the invention produce increased yields
 CC of indolocarbazoles compared with Saccharothrix aerocolonigenes
 CC ATCC39243, and may permit the production of novel indolocarbazole
 CC compounds (e.g., halogenated derivatives). The present sequence
 CC represents the specifically claimed rebeccamycin biosynthetic gene
 CC cluster from Saccharothrix aerocolonigenes ATCC39243.

Sequence 25681 BP; 4040 A; 9104 C; 8809 G; 3728 T; 0 U; 0 Other;

Qy	495	CCTGGCGCTCAACGAGAGCGGCGGAGATGCCGTGCCGAGA	554
Db	513	TCTGGTCTCACCCTCTCTCAGCGGCCCCCTCGCCCCGAGGGCGGCGCCT	572
Qy	555	CAGTG	560
Db	573	CCAGCG	578

Search completed: August 29, 2005, 15:06:02
 Job time : 593 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 29, 2005, 14:40:45 ; Search time 3416 Seconds
(without alignments)
9226.351 Million cell updates/sec

Title: US-10-069-353a-7

Perfect score: 828
Sequence: 1 gctgtccaggtgctgaccc.....cggcgaagcggctgctcga 828

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_esc1:*
2: gb_esc2:*
3: gb_hcc:*
4: gb_esc3:*
5: gb_esc4:*
6: gb_esc5:*
7: gb_esc6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88.4	10.7	1016	7	CK207033 FGAS01864
2	87.8	10.6	695	5	BQ838139 WHE2907_A
3	84.4	10.2	787	6	CA764573 AP53-RP1
4	82.6	10.0	587	7	CN008962 WHE2647_B
5	80.6	9.7	878	7	CK196136 FGAS00458
6	80.4	9.7	492	1	AV624709 AV624709
7	80.4	9.7	680	4	BG858904 1024060C0
8	80.4	9.7	730	4	BG861169 1024076A0
9	80.4	9.7	803	4	BG847555 1024017H1
10	79.8	9.6	571	2	BF257939 HVSMEF001
11	79.8	9.6	618	6	CB868921 HCO9M21W
12	79.8	9.6	622	6	CA210798 SCEPSR113
13	77.6	9.4	666	6	CD442528 EL01N0412
14	77.6	9.4	874	9	CC717854 OGTA10TM
15	77.6	9.4	908	9	CC717850 OGTA10TC
16	76.8	9.3	912	9	CC720063 OGUER12TV
17	76.8	9.3	427	5	BQ823320 1030108C1
18	74.4	9.0	554	9	CC623301 OGUHV66TH
19	74.4	9.0	658	9	CC685816 OGRHUS87V
20	74.2	9.0	1113	7	CK217823 FGAS02982
21	74	8.9	813	7	CF884047 t1c037xj
22	71.6	8.6	610	4	BF531904 1024117D1
23	71.6	8.6	653	4	BF531362 1024113C0
24	70.6	8.5	738	7	CO131978 GR_Eb441

25	70.6	8.5	799	7	CO118507 GR_Eb020
26	70.6	8.5	855	7	CO076561 GR_Ea37P
27	70.6	8.5	866	7	CO116523 GR_Eb018
28	69.8	8.4	705	7	CO435133 7d5-c3 Mo
29	69	8.3	794	7	CO091150 GR_Ea11H
30	68.8	8.3	588	2	BF619597 HVSMEC000
31	68.8	8.3	719	5	BQ838223 WHE2907_H
32	68.8	8.3	781	6	CA405371 9 A. brax
33	67.2	8.1	814	7	CN812339 Fg13 0192
34	66.8	8.1	414	1	AV638404 AV638404
35	65.8	7.9	714	6	CB970276 CAB10003
36	65.6	7.9	742	7	CO524520 3530_1_16
37	65.4	7.9	740	2	BE602586 HVSMEH010
38	64.6	7.8	696	7	CO435474 11f12_P12
39	64.6	7.8	699	7	CO435049 6c7 CAS-0
40	64.4	7.8	311	6	C91811 C91811 Rice
41	64.4	7.8	541	5	BQ134030 sams6d04
42	63.6	7.7	777	7	CO127869 GR_Eb12K
43	63.6	7.7	780	7	CO124614 GR_Eb07K
44	63.6	7.7	792	7	CO131974 GR_Eb441
45	63.6	7.7	802	7	CO123977 GR_Eb061

ALIGNMENTS

RESULT 1
LOCUS CK207033
DEFINITION FGAS018648 Triticum aestivum FGAS: Library 5 GATE 7 Triticum
aestivum cDNA, mRNA sequence.
ACCESSION CK207033
VERSION CK207033.1 GI:39569423
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 1016)
Allard, F., Crosby, W.L., Danyluk, J., Budes, F., Frick, M., Gaudet, D.,
Genewein, B., Graf, R., Gulick, P., Hrycan, L.D., Iaroché, A.,
Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilsson, D.,
Penikar, C., Roach, J.L. and Sarhan, F.
Functional Genomics of Abiotic Stress in Wheat and Canola Crops
Unpublished (2003)
Contact: Wm L Crosby
Bioinformatics
University of Saskatchewan, Department of Computer Science
1C101 Engineering Building, 57 Campus Drive, Saskatoon,
Saskatchewan, S7N 5A9, Canada
Tel: 306 966 1769
Fax: 306 966 2033
Email: fgas_esc@cs.usask.ca
This sequence is the direct result of the Base calling software
phred (default parameters). It is the raw base calls. To aid in the
identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [31,821].
Plate: L5B009 row: G column: 22.
Location/Qualifiers
1..1016
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone_lib="Triticum aestivum FGAS: Library 5 GATE 7"
/note="Vector: PCMV.SPORT6; Crown and developmental stages
of spike formation in wheat cultivar Norstar. 4 mRNA
populations were combined before constructing the library.
The first mRNA population is from 1cm crown sections after
30 days of cold acclimation. The second is from 1cm crown
sections after 11 days of deacclimation (before
deacclimation plants were fully vernalized for 49 days).

FEATURES

source

The third is from different developmental stages of spike formation (5 to 50mm) that still have not emerged from the leaf (dissection required). The last is from different developmental stages of spike and seed formation after having emerged from the leaf (visible). First strand synthesis in this library was done in the presence of methylated dCTP thereby protecting from internal cleavage with NotI."

ORIGIN

Query Match 10.7%; Score 88.4; DB 7; Length 1016;
Best Local Similarity 52.0%; Pred. No. 1.2e-10;
Matches 197; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

139 CAGCAGCGCGCGGCTCAACGACCTTGTGCGCAACGAGCGGTCTGATGGCGC 198
407 CGGAGCGGACCGGGGTGACGAGGAGCGCGGCTCTGAGGCGGCGCGG 466
199 GTTCACTGCTGATGTGGGTGCGGTACCGGACCAACGCGCTGCGCGCGAC 258
467 AAGCGGCTGCTGACGTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGAC 526
259 AACGGATTCAGATCAACCGGCTCAACCGGACGAGTGAAGTGGCCATGGCGCTCAT 318
527 TCCGAGCTCGAGCTGTGCGGATCAACGATCAAGTACAGTCAACCGCGCGCTCA 586
319 TCGCAGCGGACCGGACTAAGCAGCGGCTGACTTCTGTGCGTGAATCCATGTC 378
587 CACAACCGGACCGGCGCTGACCGGCAATGCGAGTGTGTGGGCAATTCATGGCC 646
379 CTGCGGTACCGGCAATGCTTTGCAACGCGGCTGTGGCGGCTGTGGAGATG 438
647 ATGCGGTTCGCGGACCGGCTTTCAGCGGCGCTACTCATGAGGCCACCTGCACG 706
439 TCCGACCGGACCGGCTCATCGGGAATCTTTCAGATCAACCGGCGGCTGCTC 498
707 CCCAGGCTGACGAGAGCTTCAACGAGGTGTCGCGGCTCAACGAGGCGGCTTAC 766
499 GCGGTACCGGAGGTGCTCA 517
767 GTCTCTTACGATGGGTGA 785

RESULT 2 695 bp mRNA linear EST 08-AUG-2002
B0838139
LOCUS B0838139
DEFINITION WHE2907_A04_B07ZS wheat aluminum-stressed root tip cDNA library
Triticum aestivum cDNA clone WHE2907_A04_B07, mRNA sequence.

ACCESSION B0838139
VERSION B0838139.1 GI:22142457
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum

REFERENCE 1 (bases 1 to 695)
AUTHORS Anderson,O.D., Chao,S., Chin,A., Close,T.J., Gustafson,J.P.,
Lazo,G.R., Rausch,C.J., Ross,K., Seaton,C.L. and Wilson,C.

TITLE The structure and function of the expressed portion of the wheat
JOURNAL genomes - Aluminum-stressed root tip cDNA library
COMMENT Unpublished (2001)

US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818

EMAIL: candersen@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
Seq primer: SK primer.
Quality sequence with phred score less than 20
Location/Qualifiers

source

1. 695
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="B081146"
/db_xref="taxon:4565"
/clone="WHE2907_A04_B07"
/tissue_type="Root tip at 1.0 to 1.5 mm stage"
/dev_stage="Seedling"
/lab_host="E. coli SOLR"
/clone_lib="Wheat aluminum-stressed root tip cDNA library"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid.
Site_1: EcoRI; Site_2: XhoI; Plants were grown under
hydroponic conditions, root tips were excised and snap
frozen, total RNA was prepared at University of
Missouri (Ross, Gustafson). Poly(A) RNA was purified, a
cDNA library was made, and the cDNA clones were in vivo
excised to give phagescript SK- phagemids in the T7 Close
lab (Chin and Close) at the University of California,
Riverside. Plasmid DNA preparations and DNA sequencing
were performed in the OD Anderson lab (all other
authors)."

ORIGIN

Query Match 10.6%; Score 87.8; DB 5; Length 695;
Best Local Similarity 52.0%; Pred. No. 1.6e-10;
Matches 197; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

139 CAGCAGCGCGCGGCTCAACGACCTTGTGCGGACGAGCGGTCTGATGGCGC 198
261 CGGAGCGGACCGGGGTGACGAGGAGCGCGGCTCTGAGGCGGCGGCGG 320
199 GTTCACTGCTGATGTGGGTGCGGTACCGGACCAACGCGCTGCGCGCGAC 258
321 AAGGCGCTGCTGACGTGTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGAC 380
259 AACGGATTCAGATCAACCGGCTCAACCGGACGAGTGAAGTGGCCATGGCGCTCAT 318
381 TCCGAGCTCGAGCTGTGCGGATTCATCAACGAGTGAACCGGCGGCTGCG 440
319 TCGCAGCGGACCGGACTAAGCAGCGGCTGACTTCTGTGCGTGAATCCATGTC 378
441 CACAACCGGACCGGCGCTGACCGGCAATGCGAGTGTGTGGGCAATTCATGGCC 500
379 CTGCGGTACCGGCAATGCTTTGCAACGCGGCTGTGGCGGCTGTGGAGATG 438
501 ATGCGGTTCGCGGACCGGCTTTCAGCGGCGCTACTCATGAGGCCACCTGCACG 560
439 TCCGACCGGACCGGCTCATCGGGAATCTTTCAGATCAACCGGCGGCTGCTC 498
561 CCCAGGCTGACGAGAGCTTCAACGAGGTGTCGCGGCTCAACGAGGCGGCTTAC 620
499 GCGGTACCGGAGGTGCTCA 517
621 GTCTCTTACGATGGGTGA 639

RESULT 3 787 bp mRNA linear EST 08-JAN-2003
CA764573
LOCUS CA764573

DEFINITION AF53-RP6_04_B21.T7.081.ab1 IRR1 Drought Stress Panicle Library
Oryza sativa (indica cultivar-group) cDNA clone C0001197 5' similar
to Delta(24)-sterol C-methyltransferase (BC 2.1.1.41), mRNA
sequence.

ACCESSION CA764573
VERSION CA764573.2 GI:27546528
KEYWORDS EST.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)

REFERENCE 1 (bases 1 to 787)
AUTHORS Bennett,J., Arumugam,K., Lafitte,R., Wen,J., Rudd,S. and
Ehhardtideae; Oryzaeae; Oryza.

TITLE Bruekiewicz, R.M.
JOURNAL IRRI Drought Stress Panicle cDNA Library
COMMENT Unpublished (2002)
On Dec 2, 2002 this sequence version replaced gi:25993828.
Contact: Richard Bruekiewicz
Biometrics and Bioinformatics Unit
International Rice Research Institute
DAPO 7777, Metro Manila, Philippines
Tel: +63-2-845-0563
Fax: +63-2-845-0606
Email: r.bruekiewicz@cgiar.org
International Rice Information System (IRIS)
http://www.iris.irri.org/ D0201196
Assignment of putative function to the sequence by S. Rudd of the
Munich Information Center for Protein Sequences
(http://mips.gsf.de)
Plate: 04 row: B column: 21.
Location/Qualifiers

FEATURES
source 1..787
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/cultivar="IR64"
/db_xref="taxon:39946"
/clone="C0001197"
/issue_type="Panicles"
/dev_stage="Flowering"
/clone_lib="IRRI Drought Stress Panicle Library"
/note="Vector: pBluescript II SK+; Water stress was applied by not watering for 4 consecutive days. Panicles were collected from control (well watered) and stressed plants at 2 days before heading, at heading, 50% flowering and 4 days after 50% flowering."

ORIGIN

Query Match 10.2%; Score 84.4; DB 6; Length 787;
Best Local Similarity 51.0%; Pred. No. 1.1e-09;
Matches 199; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 128 GGGCTTCTGGGAGGAGGCGGCGGCTCAACCACTTGTGCGGAAAGGACCGTGC 187
DB 121 GAGAGTCTTCTATCAAGGCGCATGTCTGCCATAGCACTACCTGCGCTCAAGATTGGCA 180
QY 188 TCGATGCGCGCGCTTGCAGCTGCTGATGTGGGGTGGCGGCTACCGGCAACCGCGCTCGGG 247
DB 181 TCAAGAGGCGCATGAAGGTTCTGATGTGGGCTGGCGGTGGTGTGCTCAAGCTCGAGAGA 240
QY 248 TCGCGGCGGCAACCGGATCCAGATCAACCGGATCAACCGTCAAGGAGTGAAGTGGCA 307
DB 241 TTGCCAAGTTCAACCGGCGGCAATTAACCGGCTCAACAACAAGATCAATCAATCAAG 300
QY 308 TCGCGGCTGATTGGCGGAGCGGAAAGCGGACTAAGCAACCGGAGTGTCTCGTGGTGC 367
DB 301 GTGCTACCCACTACGCGGCAAGAACAGAGAGCTCTCTCAACGATTTGTCAAGGGCG 360
QY 368 ATGCCATGTCCTCCGCGGTCACCGGCAATGCTTTTCAACCGCGGCTGGGCGCATGCGC 427
DB 361 ACTTCATGACAGATGCTCTTCTCGACAACTCTTCAGCGGTGTAACGCCATGAGGCTTA 420
QY 428 TGTGAGAGATGTCCGAACCGGACCGTGCATCCGCGGAAATCTTTCAGATCTCAAAACCG 487
DB 421 CCTGCCACGCGCCCAACCTCGAGAGGCAATTAACGCGAGATCTTCGTGTCTCAAGCCG 480
QY 488 GTGGCAATCTCGGCGCTACCGAGTGTCTCA 517
DB 481 GCGGCGTCTTCGGTGTATTACAGATGAGCTCA 510

RESULT 4
CN008962 587 bp mRNA linear EST 29-MAR-2004
LOCUS CN008962
DEFINITION WHE2647_B12_D23Z Wheat Fusarium graminearum infected spike cDNA
library Trifolium aestivum cDNA clone WHE2647_B12_D23, mRNA
sequence.

ACCESSION CN008962
VERSION CN008962.1 GI:45798994
KEYWORDS EST
SOURCE Trifolium aestivum (bread wheat)
ORGANISM Trifolium aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 587)

REFERENCE
AUTHORS Anderson, O.D., Chao, S., Han, P.S., Heinen, S., Heia, C.C., Kang, Y.,
Krueger, W.M., Lazo, G.R., Miller, S., Muehlbauer, G.J., Miller, R.,
Pritsch, C., Rausch, C.J., Seaton, C.L., Tong, J.C., Vance, C. and
Wilson, C.F.

TITLE The structure and function of the expressed portion of the wheat
genomes - Fusarium graminearum infected spike cDNA library
JOURNAL Unpublished (2001)
CONTACT: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov

Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20. No effort was taken
to identify ESTs of fungal origin from this library, thus this EST
could be of wheat or fungal origin.
Seq primer: 73 primer.

FEATURES
source 1..587
Location/Qualifiers

/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Sumai3"
/db_xref="taxon:4565"
/clone="WHE2647_B12_D23"
/issue_type="Spike"
/dev_stage="Adult plant"
/lab_host="E. coli SDR"
/clone_lib="Wheat Fusarium graminearum infected spike cDNA
library"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid
pBluescript SK, Site 1: EcoRI, Site 2: XhoI; Plants were
grown in the greenhouse. Spikes were sprayed at anthesis
with Fusarium graminearum. Total RNA, and poly(A) RNA were
prepared and pooled from infected spike at 0, 6, 12, 24,
36 and 48 hours after inoculation, a cDNA library was
made, and the cDNA clones were in vivo excised to give
pBluescript phagemids in G. Muehlbauer lab at the
University of Minnesota (Krueger, W.M., Muehlbauer, G.J.,
Pritsch, C., Vance, C.). The cDNA library should contain
genes of both wheat and fungal pathogen origin. Plasmid
DNA preparations and DNA sequencing were performed in the
OD Anderson lab (all other authors)."

ORIGIN

Query Match 10.0%; Score 82.6; DB 7; Length 587;
Best Local Similarity 50.4%; Pred. No. 2.9e-09;
Matches 202; Conservative 0; Mismatches 199; Indels 0; Gaps 0;

QY 134 CCTGGACAGAGCGCGCGGACCGGCTCAACCACTTGTCCGCGAAAGGACCGTGTGATG 193
DB 15 CTTTATCAAGGCGCATTTGCGGTCATGAGCACTACCTTCCCAACGATCGGATCAAG 74
QY 194 GCGGCGTTCGACTGCTCGATGTGGGTGGGTGAACCGAACAACGAGCGCTGGCGTGGCC 253
DB 75 ACGGCAATGAAGGTTCTTGATGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 134
QY 254 GCGACAGCGCATTCAGATCAACCGGATCAACCGGATCAACCGGATGAGTGGCATGCGCG 313
DB 135 AGTTCAAGAGTGTCTATTTACCGGCTTAACAACAACAATCAAGATGAGGAGGCTTA 194
QY 314 CTGATTGCGGACCGGAAACCGGACTAAGCCACCGGATGACTTCTCGTGTGATGCCA 373

Db 195 CCCACTGACGCTTCAAGAGAGGCTGTGTCCAACCACTTGATTGTTCAGAGGTGATTCA 254
 QY 374 TGTCCCTGCGCTACCGGCAATGCTTTTGAGCGCGGCTGAGGCGCAATGAGTGTGTTGG 433
 Db 255 TGCAGATGCTTTTCCCGACACAGCTTGATGATCGGTTATGCAATTGAGCGACCGTTTC 314
 QY 434 AGATGTCCGAAACCGGACCGTGCATCCGGAAATCTTCAGTACTCAACCCCGTGGCA 493
 Db 315 AGGCTCCGACTCTCAAGGGCATCTACAGGAGATTTTCGTGTTCTTAAAGCCGGTGGG 374
 QY 494 TCTCTGGGCTCACCGAGTGTCAACGAAAGCGCGGCGC 534
 Db 375 TCTCGGTGTGTACGAGTGTCTCATGACTGACGAGTACGAC 415
 RESULT 5
 CK196136 878 bp mRNA linear EST 08-DEC-2003
 LOCUS FGAS004583 Triticum aestivum FGAS: Library 3 Gate 6 Triticum
 DEFINITION aestivum cDNA, mRNA sequence.
 ACCESSION CK196136
 VERSION CK196136.1 GI:39558526
 KEYWORDS EST
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae; Triticaceae; Triticum.
 1 (bases 1 to 878)
 Allard, F., Crosby, W. L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D., Geneswein, B., Grat, R., Gulick, P., Hrycan, L. D., Laroche, A., Links, M. G., McCarthy, E. L., Monroy, A., Muzak, I., Nilsson, D., Peniket, C., Roach, J. L. and Sarhan, F.
 Functional Genomics of Abiotic Stresses In Wheat and Canola Crops
 Unpublished (2003)
 CONTACT: Wm L Crosby
 Bioinformatics
 University of Saskatchewan, Department of Computer Science
 1C101 Engineering Building, 57 Campus Drive, Saskatoon,
 Saskatchewan, S7N 5A9, Canada
 Tel: 306 966 1769
 Fax: 306 966 2033
 Email: fgas_estes@cs.usask.ca
 This sequence is the direct result of the Base calling software Phred (default parameters). It is the raw base calls. To aid in the identification of the high quality insert the software Lucy (default parameters) has been run on this sequence. Lucy identified the region [102,750].
 Plate: L3C105 row: O column: 10.
 Location/Qualifiers
 1..878
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /db_xref="taxon:4565"
 /clone_lib="Triticum aestivum FGAS: Library 3 Gate 6"
 /note="Organ: Root; Vector: PCWV.SPORT6; Root tissue from control, cold-acclimated and salt stressed wheat cultivar Norstar. 7 mRNA populations were combined before constructing the library: 7 day non-acclimated roots, 1, 23, and 53 days cold-acclimated at 4C, and 30 minutes, 3 hours and 6 hours treated roots with 200mM NaCl. Non-acclimated and cold-acclimated plants were grown in vermiculite while salt stressed plant were grown hydroponically. First strand synthesis in this library was done in the presence of methylated dNTP thereby protecting from internal cleavage with NotI."

ORIGIN

Query Match 9.7%; Score 80.6; DB 7; Length 878;
 Best Local Similarity 51.4%; Pred. No. 8.9e-09;
 Matches 185; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

QY 133 TCTGCGACGACGCGCGGCTCACCGACTTGTGCGGAAAGGACCGTGTGAT 192

Db 499 TCCACGCGACGCGCACCGCGGTCCACGAGGAGCGCGTCCGCACTGTGGGGGCAAG 558
 QY 193 GCGCGCTTCCACCTGCTGATGTGGGTGCGGTACCGGACAAACGAGCGGTCCGGTGGC 252
 Db 559 CCGGCGCACCGGTGCTGACGTGGCTGCGGCTCGCGGGCCCATATGCGGCATGCGG 618
 QY 253 CGGCAACAGGATTCAGATCACCGGATCACCGTCAAGCCAGTGAAGTGGCCATCGCC 312
 Db 619 GCGACTCCGGCGGCGCGGTGTGTGGCATCACATCAACAGATACAGGTAAACGGGGC 678
 QY 313 GGTATGTGCGACGCGGACGCGGACTAAGCCACCGGGTGAATTCTGTGCTGTGATGCC 372
 Db 679 CGCGCGCAACCGCAAGGGCGGGCTGAGCCGCGANTGCGAGGTGTGTGCGCAACTTC 738
 QY 373 ATGTCCCTGCGGTACCGGCAATGCTTTGAGCGCGGCTGGGCAAGTGTGCTTGG 432
 Db 739 ATGGCATGCTCCCTTGTGACGACGCTCTTCAAGCGGCTACTCATGAGGCCACTTGC 798
 QY 433 GAGATGTCCGAACCGGACCGGTGCATCCGGAAATCTTCGAGTACTCAACCCGTTGCG 492
 Db 799 CACGCGCCACGCTGACGAGAGTGTACGGGAGGTGTACCGGCTGTCTATGCGGGCGC 858

RESULT 6
 AV624709 492 bp mRNA linear EST 15-DEC-2000
 LOCUS AV624709 Chlamydomonas reinhardtii 5% to 0.04% CO2 Chlamydomonas
 DEFINITION reinhardtii cDNA clone LC081e06_r 5', mRNA sequence.
 ACCESSION AV624709
 VERSION AV624709.1 GI:10773886
 KEYWORDS EST
 SOURCE Chlamydomonas reinhardtii
 ORGANISM Chlamydomonas reinhardtii
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadales; Chlamydomonas.

REFERENCE
 AUTHORS 1 (bases 1 to 492)
 Asamizu, E., Miura, K., Kuchino, K., Inoue, Y., Fukuzawa, H., Ohnaka, K., Nakamura, Y. and Tabata, S.
 Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii
 DNA Res. 7 (5), 305-307 (2000)
 JOURNAL MEDLINE
 PUBMED 11089912
 CONTACT: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
 Location/Qualifiers
 1..492
 /organism="Chlamydomonas reinhardtii"
 /mol_type="mRNA"
 /strain="C9"
 /db_xref="taxon:3055"
 /clone_lib="LC081e06_r"
 /note="Vector: pBluescriptII SK-; Site 1: EcoRI; site 2: XhoI; The cDNA library was constructed from cells cultured in a carbon stress acclimatized condition in which carbon dioxide concentration in the bubbling gas was changed from 5% to 0.04%"

FEATURES

source

ORIGIN

Query Match 9.7%; Score 80.4; DB 1; Length 492;
 Best Local Similarity 51.1%; Pred. No. 9.6e-09;
 Matches 189; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

QY 146 CCGCGCACCGGCTACCGACTTGTGCGCGGAAACCGGCTGTGATGCGCGGCTTGAC 205
 Db 53 CCGAGCGCGCGGACGAGGCGGATCTGCTGCGGCTGTGACCGCGGCAAGG 112
 QY 206 TGTCTGATGTGGGTGTGGTACCGGACCAACGCGTGGCGGCGGCAACCGCA 265

```

Db      113 CGGTGACATCGCGCTCGGCTGCGGTGGCCCATGGCTACCGTCTGCTGCAGGGGG 172
QY      266 TCCAGATCAACCGGATCACCCTGACGCCAGGTGCAAGTGGCCATCGCCGCTATTGGCCAC 325
Db      173 CCGACATCAACCGGATCACCCTGACGCCAGGTGCAAGTGGCCATCGCCGCTATTGGCCAC 325
QY      326 GCGAAGCGGATCAACCGGATCACCCTGACGCCAGGTGCAAGTGGCCATCGCCGCTATTGGCCAC 385
Db      233 CACGGAAGGTGCTGCGCTCGGCTGACGAGCTTGTCCGCGGCACTTCACTAACATCCGT 292
QY      386 ACCCGGACATGCTTTCGACGCGCGCTGGGCAATGCAAGTGTGAGATGTCGAAC 445
Db      293 TCAAGAGAAACACCTTCGACGCGGCTTACGCTTTAGAGGCACTTCCACGCCCAAGC 352
QY      446 CCGACCGTGCATCCGGAATCTTCGAGTACTCAACCCGGTGCATCTTCGCGCTCA 505
Db      353 TGGAGCAGGTGTACGCGGAGATCTACCGGCTCAAGCCCGGCACTTCTGTGTCT 412
QY      506 CCGAGTGTCT 515
Db      413 ACGAGTGGT 422

```

```

RESULT 7
BG858904      680 bp mRNA linear EST 29-MAY-2001
LOCUS      1024060C08.y1 C. reinhardtii CC-1690, normalized, lambda Zap II
DEFINITION      Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION      BG858904
VERSION      BG858904.1 GI:14240088
KEYWORDS      EST.
SOURCE      Chlamydomonas reinhardtii
ORGANISM      Chlamydomonas reinhardtii
REFERENCE      1 (bases 1 to 680)
AUTHORS      Grossman, A., Davies, J., Federpspiel, N., Harris, E., Lefebvre, P.,
            McDermott, J.P., Silflow, C., Stern, D., and Surzycki, R.
TITLE      Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants; project phase 2
JOURNAL      Unpublished (2000)
COMMENT      Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu.
FEATURES
    source
        1..680
            /organism="Chlamydomonas reinhardtii"
            /mol_type="mRNA"
            /strain="CC-1690 wild type mt+ 21gr"
            /db_xref="taxon:3055"
            /clone_1ib="C. reinhardtii CC-1690, normalized, lambda Zap
            II"
            /note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
            XhoI; This library, constructed by John Davies and Jeffrey
            McDermott, combines cDNAs from CC-1690 cells grown to
            mid-log phase in YAP (acetate-containing) medium in the
            light, YAP medium in the dark, HS (minimal) medium in
            ambient levels of CO2 and HS medium bubbled with 5% CO2.
            PolyA mRNA was purified from each sample, pooled and cDNA
            synthesized. The cDNA was directionally cloned into lambda
            Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
            pBluescript II SK- plasmids were excised from the lambda
            Zap clones by superinfection with Exsist (Stratagene)
            phage. The library was normalized using method 4 described
            in Bonaldo et al (1996) Genome Research 6: 791-806."

```

ORIGIN

```

Query Match      9.7%; Score 80.4; DB 4; Length 680;
Best Local Similarity 51.1%; Pred. No. 9.8e-09;
Matches 189; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

QY      146 CCGCGGACCGGCTCACCGACCTTGTGCGCGAAAGGACCGTGTCCATGCGCGCTTCGAC 205
Db      224 CCGAGGCGCGGACGAGGCGCGCATCGCTCTCGCTGCGCTGCAAGCCGCGCGAGAAAG 283
QY      206 TCGTGAATGTGGGGGCGGTACCGGACAAACGAGCGCTGCGCGCGGACAAAGGCA 265
Db      284 CGGTGACATGCGCGCTGCGGTGTGGGGCGCGCATCGCTGCTGTGTGACGCGG 343
QY      266 TCCAGATCAACCGGATCACCCTGACGCCAGGTGCAAGTGGCCATCGCCGCTATTGGCCAC 325
Db      344 CCGACATCAACCGGATCACCCTGACGCCAGGTGCAAGTGGCCATCGCCGCTATTGGCCAC 403
QY      326 GCGAAGCGGATCAACCGGATCACCCTGACGCCAGGTGCAAGTGGCCATCGCCGCTATTGGCCAC 385
Db      404 CACGGAAGGTGCTGCGCTCGGCTGACGAGCTTGTCCGCGGCACTTCACTAACATCCGT 463
QY      386 ACCCGGACATGCTTTCGACGCGCGCTGGGCAATGCAAGTGTGAGATGTCGAAC 445
Db      464 TCAAGAGAAACACCTTCGACGCGGCTTACGCTTTAGAGGCACTTCCACGCCCAAGC 523
QY      446 CCGACCGTGCATCCGGAATCTTCGAGTACTCAACCCGGTGCATCTTCGCGCTCA 505
Db      524 TGGAGCAGGTGTACCGGAGATCTACCGGCTCAAGCCCGGCACTTCTGTGTCT 583
QY      506 CCGAGTGTCT 515
Db      584 ACGAGTGGT 593

```

```

RESULT 8
BG861169      730 bp mRNA linear EST 29-MAY-2001
LOCUS      1024076A09.y1 C. reinhardtii CC-1690, normalized, lambda Zap II
DEFINITION      Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION      BG861169
VERSION      BG861169.1 GI:14242353
KEYWORDS      EST.
SOURCE      Chlamydomonas reinhardtii
ORGANISM      Chlamydomonas reinhardtii
REFERENCE      1 (bases 1 to 730)
AUTHORS      Grossman, A., Davies, J., Federpspiel, N., Harris, E., Lefebvre, P.,
            McDermott, J.P., Silflow, C., Stern, D., and Surzycki, R.
TITLE      Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants; project phase 2
JOURNAL      Unpublished (2000)
COMMENT      Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu.
FEATURES
    source
        1..730
            /organism="Chlamydomonas reinhardtii"
            /mol_type="mRNA"
            /strain="CC-1690 wild type mt+ 21gr"
            /db_xref="taxon:3055"
            /clone_1ib="C. reinhardtii CC-1690, normalized, lambda Zap
            II"
            /note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
            XhoI; This library, constructed by John Davies and Jeffrey
            McDermott, combines cDNAs from CC-1690 cells grown to
            mid-log phase in YAP (acetate-containing) medium in the
            light, YAP medium in the dark, HS (minimal) medium in
            ambient levels of CO2 and HS medium bubbled with 5% CO2.

```

PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with Exsist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."

ORIGIN

Query Match 9.7%; Score 80.4; DB 4; Length 730;
Best Local Similarity 51.1%; Pred. No. 9.8e-09;
Matches 189; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

146 CCGCCGACCGGCTCACCGACCTTGTCCGACGACCGTGTGATGGCGGCTTGCAC 205
|||
10 CCGAGGCGCGGACGAGGCGCGCATGCTGCTGCTGCGCTGACGCCGCGCAGAAAG 69
|||
206 TGCTCGATGTGGGGTGGGTACCGGACCAACGACGCTGCGCGCGACCAACGCGA 265
|||
70 CGCTGGACTGCGGCTGGGTGGGTGGGCGCGATCGCTGCTGCTGCTGCTGCTGCTG 129
|||
266 TCACATATCACCGGATCACCTGACGACGACGACGACGACGACGACGACGACGAC 325
|||
130 CGCATATCACCGGATCACCTGACGACGACGACGACGACGACGACGACGACGACG 189
|||
326 GCGAAGCGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 385
|||
190 CACGCGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 249
|||
366 ACCCGACATGCTTTCGACGCGCGCTGGCGCATGCTGCTGCTGCTGCTGCTGCTG 445
|||
250 TCAAGAGAGAACCTTTCGACGCGCGCTGGCGCATGCTGCTGCTGCTGCTGCTGCTG 309
|||
446 CCGACCGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 505
|||
310 TGGAGCAGGTGTACGCGGACGATCTACCGGCTGCTCAAGCGCGACGATCTGCTGCT 369
|||
506 CCGAGGTGCT 515
|||
370 ACGAGTGGGT 379

RESULT 9 803 bp mRNA linear EST 29-MAY-2001
BG847555
LOCUS 1024017H12.y1 C. reinhardtii CC-1690, normalized, lambda Zap II
DEFINITION Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION BG847555
VERSION BG847555.1 GI:14228739
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Volvocales;
Chlamydomonadales; Chlamydomonas.

REFERENCE 1 (bases 1 to 803)
Grosman,A., Davies,J., Federpiel,N., Harris,E., Lefebvre,P.,
McDermott,J.P., Silflow,C., Stern,D. and Surzycki,R.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants, project phase 2
Unpublished (2000)

JOURNAL COMMENT
Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.

FEATURES
source 1. 803
Location/Qualifiers

/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21g"

/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, normalized, lambda Zap
II"

/note="Vector: pBluescript II SK-, Site 1: EcoRI, Site 2:
XhoI; This library, constructed by John Davies and Jeffrey
McDermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in YAP (acetate-containing) medium in the
light, YAP medium in the dark, HS (minimal) medium in
ambient levels of CO2 and HS medium bubbled with 5% CO2.
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
pBluescript II SK- plasmids were excised from the lambda
Zap clones by superinfection with Exsist (Stratagene)
phage. The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."

ORIGIN

Query Match 9.7%; Score 80.4; DB 4; Length 803;
Best Local Similarity 51.1%; Pred. No. 9.9e-09;
Matches 189; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

146 CCGCCGACCGGCTCACCGACCTTGTCCGACGACCGTGTGATGGCGGCTTGCAC 205
|||
10 CCGAGGCGCGGACGAGGCGCGCATGCTGCTGCTGCGCTGACGCCGCGCAGAAAG 69
|||
206 TGCTCGATGTGGGGTGGGTACCGGACCAACGACGCTGCGCGCGACCAACGCGA 265
|||
70 CGCTGGACTGCGGCTGGGTGGGTGGGCGCGATCGCTGCTGCTGCTGCTGCTGCTG 129
|||
266 TCACATATCACCGGATCACCTGACGACGACGACGACGACGACGACGACGACGAC 325
|||
130 CGCATATCACCGGATCACCTGACGACGACGACGACGACGACGACGACGACGACG 189
|||
326 GCGAAGCGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 385
|||
190 CACGCGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 249
|||
366 ACCCGACATGCTTTCGACGCGCGCTGGCGCATGCTGCTGCTGCTGCTGCTGCTG 445
|||
250 TCAAGAGAGAACCTTTCGACGCGCGCTGGCGCATGCTGCTGCTGCTGCTGCTGCTG 309
|||
446 CCGACCGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 505
|||
310 TGGAGCAGGTGTACGCGGACGATCTACCGGCTGCTCAAGCGCGACGATCTGCTGCT 369
|||
506 CCGAGGTGCT 515
|||
370 ACGAGTGGGT 379

RESULT 10 571 bp mRNA linear EST 22-OCT-2001
BF257939
LOCUS HVSMEF0014E17f Hordeum vulgare seedling root EST library HVCDA0007
DEFINITION (Ectiolated and unstressed) Hordeum vulgare subsp. vulgare cDNA
clone HVSMEF0014E17f, mRNA sequence.
ACCESSION BF257939
VERSION BF257939.2 GI:13119150
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Hordeum.

REFERENCE 1 (bases 1 to 571)
Wing,R., Close,T.J., Kleinhoofs,A., Wise,R., Begum,D., Frisch,D.,
Yu,Y., Henry,D., Palmer,M., Rambo,T., Simons,J., Choi,D.W.,
Fenton,R.D., Oates,R. and Main,D.
Development of a genetically and physically anchored EST resource
for barley genomes: Morex unstressed seedling root cDNA library
unpublished (2001)

JOURNAL COMMENT
On Nov 16, 2000 this sequence version replaced gi:11187052.

Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: twing@clemson.edu
Total hg bases = 319
Seq primer: AATTAACTCACTTAAGG
High quality sequence stop: 480.
Location/Qualifiers

FEATURES

1..571
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Morex"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="HVSMEF0014E17F"
/issue_type="Seedling root"
/lab_host="TU121"
/clone_lib="Hordeum vulgare seedling root EST library
HYCDNA0007 (Etiolated and untreasured)"
/note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;
Seeds were surface sterilized then germinated under axenic
conditions in the dark at room temperature on filter paper
with water, nystatin and ceftaxime in covered
crystallization dishes. Five-day old seedling roots were
then harvested, total RNA was prepared, poly(A) RNA was
purified, one primary unamplified cDNA library was made,
and 1 million pfu were in vivo excised to give pluescript
SK(-) cDNA phagemids. These steps were performed in the TU
Close Laboratory at the University of California,
Riverside (Choi, Close, Fenton). Phagemids were plated and
picked at the Clemson University Genomics Institute (CUGI)
(Begum, Palmer, Friesch, Atkins and Wing). Plasmid DNA
preparations, DNA sequencing and sequence analysis were
performed at CUGI (Wing, Yu, Friesch, Henry, Simmons,
Oates, Rambo, Main). The sequence has been trimmed to
remove vector sequence and contains a minimum of 100 bases
of phred value 20 or above. For more details on library
preparation and sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders Also
see Close TJ, Wing R, Kleinbols A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/gspages/bgn/31/cover.html)"

ORIGIN

Query Match 9.6%; Score 79.8; DB 2; Length 571;
Best Local Similarity 51.2%; Pred. No. 1.3e-08;
Matches 186; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

139 CACGAGGCGGCGGACCGGCTCACCGACCTTGTGCGGAAGGACCGGTGCGATGGGGC 198
149 CGGAGGCGGACCGGCTGACGAGAGCGGCTTCCACCTGCTGAGGCGCGCCGGG 208
199 GTTCGACTGCTCATGTGGGGTGGCGTACCGGACCAACGACGCTGCGCGCGCGAC 258
209 AACCGGCTGCTGACGCTGCGGCTGGGGGCGGCGCCATGCGCGCATGGGGCGAC 268
259 AACGCGATTCAGATCAACCGGATCAACCGGATGCAAGTGGCCATGCGCGTAT 318
269 TCCGGGCTCCGACGCTCGGGATCACTATCAACGATACAGGTCAACCGCGCGCTCG 328
319 TGGGACGCGGACCGGACTTAAGCCACCGGGTGAAGTTCGTCGTCGTCGATGCTTC 378
329 CACGACCGCATGGCGGCTGGAACCAATGCAAGGTGCTGGGCACTTCATGGCC 388
379 CTGCGGTAACCGGACCAATGCTTTCGACGCGCGCTGGGCGATGAGTGCCTTTGAGATG 438
389 ATGCGCGTCCGACGCGCTTTCGACGCGGCGCTACTCATCAAGGCGACCTGCGACGCG 448

QY 439 TCCGACCGGACCGGCTCCATCCGGAATCCTTTCAGTACTCAACCGGTGCGATCCGC 498
DB 449 CCGAGCTTTCGACGACTTCAAGGAGGAGCTTCCAGTCTCAAGCGGCGGCTTAC 508
QY 499 GGC 501
DB 509 GTC 511

RESULT 11

CB868921

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Barley ESTs from coleoptile tissue
Unpublished (2003)
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert length: 618 Std Error: 0.00
Plate: 9 row: M column: 21
Seq primer: T7.

FEATURES

source

1..618
/organism="Hordeum vulgare"
/mol_type="mRNA"
/cultivar="Sloop"
/db_xref="GABI:547323"
/clone="HC09W21"
/issue_type="coleoptile"
/dev_stage="coleoptile, 3 days old"
/lab_host="DH10B"
/clone_lib="CH"
/notes="Vector: pSPORT; Site 1: SalI (5'-end of cDNA);
Site 2: NotI (3'-end of cDNA); Due to the cloning system
used blue/white selection for recombinants is not 100%
reliable. Average insert size is 1.3 kb."

ORIGIN

Query Match 9.6%; Score 79.8; DB 6; Length 618;
Best Local Similarity 52.9%; Pred. No. 1.4e-08;
Matches 171; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

195 CCGCGTTCGACTGCTCATGTGGGGTGGCGTACCGGACCAACGCGCTGCGCGCG 254
DB 3 CCGGAAGCGGCTCTGAGAGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGG 62
QY 255 CGACGACGCGATTCAGATCAACCGGATCAACCGGATGCAAGTGGCCATGCGCGC 314
DB 63 GCACTCGGCTCCGACGCTGCGGATCACTATCAACGATCAACGCGGCGCGG 122
QY 315 TGAATGCGACGCGGACCGGACTAAGCCACCGGGTGAAGTTCGTCGTCGATGCTTC 374
DB 123 CTGCAACGCGGACCGGCGGCTGGAACCAATGCAAGGTGCTGGGCACTTCAT 182
QY 375 GTCCGCGGTAACCGGACCAATGCTTTCGACGCGCGCTGGGCGCATGCTGTTGGA 434

Db 183 GGCATGCGCTTCCGAGACGCGCTCCTTCGACGCGCCTTACCTTCATCGAGGCCACCTGCCA 242

Qy 435 GATGTCGACGACCGGACCGGTCATCCCTTCGAGTACTCAAAACCGGATGAGAT 494

Db 243 CGCGCCCGAGCTTCGACGAGCTTACGCGGAGCTTCCGCGTCTCAAGCCAGGGGGGCT 302

Qy 495 CCTGGCGCTCACCGAGGTCTGCA 517

Db 303 CTACGCTTCATACGAGTGGGTCA 325

RESULT 12

CA210798 622 bp mRNA linear EST 25-SEP-2003

LOCUS SCEP5B1132A04.g SBI Saccharum officinarum cDNA clone SCEP5B1132A04

DEFINITION 5', mRNA sequence.

ACCESSION CA210798

VERSION CA210798.1 GI:35254474

KEYWORDS EST.

SOURCE Saccharum officinarum

ORGANISM Saccharum officinarum

REFERENCE Bukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Saccharum; Saccharum officinarum complex.

AUTHORS 1 (bases 1 to 622)

TITLE Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.

JOURNAL The libraries that made SUCEST

COMMENT Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: pattud@unicamp.br

Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at <http://www.bcccenter.fcav.unesp.br>

Plate: 132 row: A column: 04

Seq primer: T7 Promoter Primer.

Location/Qualifiers

1. 622

/organism="Saccharum officinarum"

/mol_type="mRNA"

/db_xref="taxon:4547"

/clone="SCEP5B1132A04"

/lab_host="DH10B"

/clone_1lb="SBI"

/note="Organ: Stalk Bark from adult plants; Vector: pSport1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [Stalk Bark from adult plants]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a separose C1-28 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.fad.ic.unicamp.br/public>"

ORIGIN

Query Match 9.6%; Score 79.8; DB 6; Length 622;

Best Local Similarity 52.1%; Pred. No. 1.4e-08;

Matches 202; Conservative 0; Mismatches 183; Indels 3; Gaps 1;

Qy 133 TCTTGGCAGACGCGCGCCGACCGGCTACCGACCTTGTGCGGAACGACCGTGTGAT 192

Db 137 TCCACCGCGACGCGCACGCGCGTCCACGAGAGCGGTCGCGACCTCTCGGCCACAG 196

Qy 193 GCGGCGCTTCGACTGTGTGATGTGGGTGCGGTACCGACAACGACGCTGCGCTCGG 252

Db 197 CGGGCGACCGGCTCTCTGACGCTGCGGCGGTGCGGCGCCCATGCGCCCATCGCC 256

Qy 253 CGCGACAACGGATCCAGATCACCGGATCACCGCTACGCCAGGTGCAAGTGGCCATGCC 312

Db 257 GCGACTTCGGGATCCAAAGTGTGTGTCATCACATCAACAGATACAGGTGAACCGCGCC 316

Qy 313 GCTATTGCGGACGCGGAACGGGACTAAGCCACCGGGT---GACTTCTGTGTGTGAT 369

Db 317 CGCGCGACAACCGCAAGCGCGGCTTCGATCCCCGCGGCGAGGTGCTGTGGGGAAC 376

Qy 370 GCGATGTCCTCGCGCTACCGGACCAATGCTTTTCGACCGCGCTGGCCATGACATCGCTG 429

Db 377 TTCCTTCATGATCCCTTCCTCCGAGAGCCTCCTTCGACGCGCGCTACTCATGAGCCACC 436

Qy 430 TTGAGATGTCGGAACCGGACCGTTCATCCGGAATTCCTTGAGTACTCAACCCGGT 489

Db 437 TGCCACGCGCCACGAGCTGCAAGAGCTACGCGGCGCAAGTCAACGCTGTCAAGCCGCGC 496

Qy 490 GGCATCTCGGCGCTCACCGAGGTCTGCA 517

Db 497 GGGCTTACGTTTCTACGAGTGGGTCA 524

RESULT 13

LOCUS CD442528 666 bp mRNA linear EST 03-JUN-2003

DEFINITION E101N0412G05.b Endosperm_4 Zea mays cDNA, mRNA sequence.

ACCESSION CD442528

VERSION CD442528.1 GI:31358171

KEYWORDS EST.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE Bukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.

AUTHORS 1 (bases 1 to 666)

TITLE Lai,J., Dey,N., Kim,C.S., Bharti,A.K., Rudd,S., Mayer,K.F., Larkins,B., Becraft,P. and Messing,J.

JOURNAL Characterization of the maize endosperm transcriptome and its comparison to the rice genome

COMMENT Genome Res. 14 (10), 1932-1937 (2004)

Contact: lai, jinsheng

Dr. Joachim Messing's lab

Waksman Institute, Rutgers University

190 Frelinghuysen Rd., Piscataway, NJ 08854, USA

Tel: 732-445-3801

Fax: 732-445-5735

Email: jlai@waksman.rutgers.edu

Seq primer: T3.

Location/Qualifiers

1. 666

/organism="Zea mays"

/mol_type="mRNA"

/culivar="W22"

/db_xref="taxon:4577"

/tissue_type="Endosperm_4"

/clone_1lb="Endosperm_4"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 9.4%; Score 77.6; DB 6; Length 666;

Best Local Similarity 51.8%; Pred. No. 4.6e-08;

Matches 201; Conservative 0; Mismatches 184; Indels 3; Gaps 1;

Qy 133 TCTTGGCAGACGCGCGCCGACCGGCTACCGACCTTGTGCGGACGACCGTGTGAT 192

Db 198 TCCACCGCGACGCGCACGCGGCTCCACGAGAGCGGCTCGCGACCTCTCGGCCCGC 257

Qy 193 GCGGCGCTTCGACTGTGTGATGTGGGTGCGGTACCGACAACGACGCTGCGCTCGG 252

Db 258 CGGGCGACCGGCTCTCTGACGCTGCGGCGGTGCGGCGCCCATGCGCGCATCGCC 317

Qy 253 CGCGACAACGGATCCAGATCACCGGATCACCGCTACGCCAGGTGCAAGTGGCCATGCC 312

Db 318 GCGCACTCGGGGTCGACGTCGTGGGATCACCATCAAGAGTACCAGGTGAGGCGCGC 317
QY 313 GCTGATTGGCAGCGGACGCGACTAAG---CCACCGGAGTGAATTCTTCGTGCTGCAT 369
Db 378 CCGGCGCAACCGGACGAGGCGCGCTCGACTCCCCCGCTGCGAGGTCTGTCCGGCAAC 437
QY 370 GCCATGTCCTGCGATACCGGACCAATGCTTTGACGCGCCCTGGGCAATGACGTGCTG 429
Db 438 TTCTCTCATGATGCTTCCCGAGCGGCTCTTGACGCGGCTTACTTCATCGAGGCGACC 497
QY 430 TTGAGATGTCGGAACCGGACCGTGCATCCGGAATCTTTCGATCTCAACCCGGT 489
Db 498 TCGCAGCGCCCGACGCTGACGAGCTACGCGAGGTCTACCGCGTCTCAAGCGCGGG 557
QY 490 GGCATCTCGGCGTCAACCGAGTCTCA 517
Db 558 GCGCTCTACGTCCTCTCAAGAGTGGTCA 585

RESULT 14
CC717854 874 bp DNA linear GSS 19-JUN-2003
LOCUS OGTAK10TM.ZM.0.7.1.5.KB.Zea.mays.genomic.clone.ZMMBMA0389B19,
DEFINITION genomic survey sequence.
ACCESSION CC717854
VERSION CC717854.1 GI:32122630
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 874)

REFERENCE
AUTHORS WhiteIaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
JOURNAL Other_GSSs: OGTAK10TC
COMMENT Contact: Cathy WhiteIaw
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteIaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
1..874
Location/Qualifiers

/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0389B19"
/clone_1lb="ZM.0.7.1.5.KB"
/note="Vector: pBCK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 9.4%; Score 77.6; DB 9; Length 874;
Best Local Similarity 51.8%; Pred. No. 4.7e-08;
Matches 201; Conservative 0; Mismatches 184; Indels 3; Gaps 1;
QY 133 TCCTGACAGAGCGCGCGACCGGCTGACCGACCTTGTGCGCGAAGCGAGCTGTGCAT 192
Db 452 TCCACCGGACGCGACGCGGTCTCAGAGGAGCGGTGCGACCTCTCGGCGCCGG 511
QY 193 GCGGCGCTTGACTGCTGATGTGGGCTGGGTACCGGACAAACGAGCGCTGCGCGC 252
Db 512 CCGGCGCACCGCTCTCTGACGTGGGCTGCGGCTCGCGCGGCGCCATGCGCATGCGC 571
QY 253 CCGGACAAACGATTCAGATCAACCGGATCAACGCTGACGAGTGCAGAGTGGCATGCGC 312

Db 572 GCGCACTCGGGGTCGACGTCGTGCGCATCACCATCAACAGATACAGGTGAGGCGCGC 631
QY 313 GCTGATTGGCAGCGGACGCGACTAAG---CAACCGGAGTGAATTCTTCGTGCTGCAT 369
Db 632 CCGGCGCAACCGGACGAGGCGCGCTCGACTCCCCCGCTGCGAGGTCTGTCCGGCAAC 691
QY 370 GCCATGTCCTGCGATACCGGACCAATGCTTTGACGCGCCCTGGGCAATGACGTGCTG 429
Db 692 TTCTCTCATGATGCTTCCCGAGCGGCTCTTGACGCGGCTTACTTCATCGAGGCGACC 751
QY 430 TTGAGATGTCGGAACCGGACCGTGCATCCGGAATCTTTCGATCTCAACCCGGT 489
Db 752 TCGCAGCGCCCGACGCTGACGAGCTACGCGAGGTCTACCGCGTCTCAAGCGCGGG 811
QY 490 GGCATCTCGGCGTCAACCGAGTCTCA 517
Db 812 GCGCTCTACGTCCTCTCAAGAGTGGTCA 839

RESULT 15
CC717850/c 908 bp DNA linear GSS 19-JUN-2003
LOCUS OGTAK10TC.ZM.0.7.1.5.KB.Zea.mays.genomic.clone.ZMMBMA0389B19,
DEFINITION genomic survey sequence.
ACCESSION CC717850
VERSION CC717850.1 GI:32122626
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 908)

REFERENCE
AUTHORS WhiteIaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
JOURNAL Other_GSSs: OGTAK10TM
COMMENT Contact: Cathy WhiteIaw
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteIaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
source
1..908
Location/Qualifiers

/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0389B19"
/clone_1lb="ZM.0.7.1.5.KB"
/note="Vector: pBCK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 9.4%; Score 77.6; DB 9; Length 908;
Best Local Similarity 51.8%; Pred. No. 4.7e-08;
Matches 201; Conservative 0; Mismatches 184; Indels 3; Gaps 1;
QY 133 TCCTGACAGAGCGCGCGACCGGCTGACCGACCTTGTGCGCGAAGCGAGCTGTGCAT 192
Db 738 TCCACCGGACGCGACGCGGTCTCAGAGGAGCGGTGCGACCTCTCGGCGCCGG 679
QY 193 GCGGCGCTTGACTGCTGATGTGGGCTGGGTACCGGACAAACGAGCGCTGCGCGC 252
Db 678 CCGGCGCACCGCTCTCTGACGTGGGCTGCGGCTGCGGCGGCGCCATGCGCATGCGC 619
QY 253 CCGGACAAACGATTCAGATCAACCGGATCAACGCTGACGAGTGCAGAGTGGCATGCGC 312

```
Db      618 GCGCACTGGGGGTCCGACGTCGTCCGATCACCATCAACGAGTACAGGTGAAGCGGCC 559
Oy      313 GCTGATTGGGACGCGAAACGGACTTAAG---CCACCGGGTGGACTTCTCGTGGTCGAT 369
Db      558 CGCGGCGACAAACGCAAGGCCGGCTTCGACTCCCGCGCTGGAGGTGCTTGGCGAAC 499
Oy      370 GCCATGTCCCTGCGCGTACCGGACAAATGCTTTGACGCGCGCTGGGCCATGCAATCGCTG 429
Db      498 TTCCTCTCCATGCTCCCTTCCCGACGCGTCTTGAAGGGCGCTTCCATCGAGGCCACC 439
Oy      430 TTGGAGATGTCGGAACCGGACCGTGCATCCGGGAAATCCTTCGAGTACTCAAAACCGGT 489
Db      438 TCCCAAGCGGCTGAGGAGCTTACGAGGAGTCTTACCGCGTCTCAAGCGCGGG 379
Oy      490 GGCATCTCTGGCGTCCAGAGGTGCTCA 517
Db      378 CGCCTTACGTCTCTACGAGTGGGTCA 351
```

Search completed: August 29, 2005, 17:07:58
Job time : 3425 secs

[illegible]

```

; SOFTWARE: Patentin version 3.0
; SEQ ID NO 21
; LENGTH: 846
; TYPE: DNA
; ORGANISM: Streptomyces platensis subsp. rosaceus
US-10-132-134-21

```

Query Match	20.5%	Score 170;	DB 16;	Length 846;
Best Local Similarity	59.4%	Pred. No. 4.3e-38;		
Matches 328;	Conservative 0;	Mismatches 215;	Indels 9;	Gaps 2;

OY	5	TGCGAGGTGGCGGACCAACATGCACTGACGCGAGGTTGGGCGAGATGTATGACCTTGATCCGCGCT	64
Db	35	TCCCTCTCTCCGCTCCCTCGTCCCGGAAGAGTGGACACTTATGACCCGCTCACCGCAC	94
OY	65	TGCTGAACTCGGTCCGGGCGGCGCCCTCGCATTCACACGCGTACTGGAGAACGAC	124
Db	95	TGGAACACCGAAGCGGCGGCG-----GAGCTTCACCTTGCTGATCTGGGAGCGTGCAG	148
OY	125	GCGGCGCTTCTGCG---AGCAGCGCCGCGACCGGCTCACCGACTTGTGCGGAAACGA	181
Db	149	ACAACGACACCCCGCTCGTGAAGCGGCGCGACCGGCTCACCGACGATGACGACCGCC	208
OY	182	CGGTGCTGCAATGGCGGCGCTTTCGACTGCTCGAATGTGGGTGGTGGTACCGGACCAACGACGC	241
Db	209	TGCGAATTCACACGAGGACAGCGGGATCTTCGACGTGCGTGGAGTCCGCGACGCGGCA	268
OY	242	TGCGCGCTCGCGGCGGACCAACGGAATCCAGATCACCGGACATCACCGTACGACGAGGTGCAC	301
Db	269	TGGGAAATCGGCGGCGGACCGGCGGCGCAATGTCAAGGGATGCGCATAGCAAGAACGACG	328
OY	302	TGACCATTCGCGCGCTGATTGCGCACCGGAAACGGGACTTAAGCCACCGGAGTGACTTCTCGT	361
Db	329	TGCGCGGCGGACCGGCTTCGCGCGAGGCGCGCGGCTGAGCGACCGGTTGAGTTCCGCGC	388
OY	362	GCGTCGATCGAATGTCTCCGCGCGCTACCGGACCAATGCTTTCGACGCGCGCTCGGCGCATGC	421
Db	389	AGCGCCAGCGCAATGAACTGCTCCCTTCCCGACGACTCTTTCGACGCGCGCATGCGCAATCG	448
OY	422	AGTCCGCTTGGAGAGATGTCCGAACCGGACCGGTGCATTCGCGGAATCTTTCGATACTCA	481
Db	449	AGTTCATCTTCAKATGCGCCGACCGCGGACCGGATCTTCGCGGAATTCGCGCGGCTACTGC	508
OY	482	AACCGGATGACATCTTCGCGGCTCACCGAGGTGTCAACGGAAGCGGCGCGGAGTGC	541
Db	509	GCCCGGCGGCGCGCTGATCTCAACGACTTCTTCGAGCGCGCGCCGCTCCCGCGGACGA	568
OY	542	CGGTGTCGCGGCG	553
Db	569	AGCAGCGCGCGG	580

```

RESULT 3
US-10-132-134-1
Sequence 1, Application US/10132134
Publication NO. US2003071562A1
GENERAL INFORMATION:
APPLICANT: Farnet, Chris
APPLICANT: Yang, Xianhu
APPLICANT: Staffa, Alfredo
APPLICANT: Zazopoulos, Emmanuel
TITLE OF INVENTION: POLYPEPTIDE SYNTHASE ENZYMES
FILE REFERENCE: 3012-205
CURRENT APPLICATION NUMBER: US/10/132,134
CURRENT FILING DATE: 2002-04-26
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 52101
TYPE: DNA
ORGANISM: Streptomyces platensis subsp. roseaceus
US-10-132-134-1

```

Query Match 20.5%; Score 170; DB 16; Length 52101;
Best Local Similarity 59.4%; Pred. No. 5.1e-38;
Matches 328; Conservative 0; Mismatches 215; Indels 9; Gaps 2;

QY 5 TGGCAGTGGCGGACCAACATCGACAGGTTGGGAGATGTATGACTGTGTACCGCCGT 64
DB 51268 TCCCCCTCTCCGCCCCCTCCCTCCGAGAGAGTGGGACCTCTACGACCCGCTACCGGCAC 51327
QY 65 TGTGTAATCTCGTTCGGGGGGGCGCCCTGCGCATTCACCACGGCTA CTGGAGAGACGACG 124
DB 51328 TGGACACCGAAGGGGCGGGG-----GCAGCTTCCACCTCGGCTACTGGAGCGTGCAGC 51381
QY 125 GGGGGGCTTCTTGGC---AGCAGCGCGCGGACCGGCTCAACCACTTTGTGCGGACGGA 181
DB 51382 ACNACACACACCCGCTGTGGAGCGGCGGACCGGCTCAACGATGACGACCGCC 51441
QY 182 CGGTGTGATGAGCGGCGGCTTGCATGTGTGGGGTGGCGGTACCGGACACAGCGC 241
DB 51442 TGGGATTCGACAGGAGACAGCGGCTCTTCAACGTGGGAGTCCGACCGCCCA 51501
QY 242 TGGCGCTGCGCGCGACCAACGCGATCCAGATCA CCGGATCACGTCAGCCAGTGCAG 301
DB 51502 TGGGATTCGCGCGGCGGACCGGCGCATGTCA CCGGATCAGCAGACGACAG 51561
QY 302 TGGCATTCGCGCGCTGATTTGGGACAGCGGACGCGGACTAAGCCACCGGTTGGAATTCTCT 361
DB 51562 TCGCCCGCGCACCGGCTCTGCGAGGGCGCGGCTGACGACCGCGTGAATTTCCGCGC 51621
QY 362 GGGTGTGATGCGATGCTCCCTGCGCGGACCGGACATGCTTTGACGCGCGCTGGGACATGC 421
DB 51622 ACCCGACCGCATGAGACTGCTTCCCGACGATCTTTGACGCGCGCATGCGCATG 51681
QY 422 AGTGTGTTGAGATGTCCGAA CCGGACCGTSCCATTCGGGAAATCTTTCAGTACTCA 481
DB 51682 AGTGTATCTTCCACATGCCGACCGGAGCGGCTCTCGCGGATCCGCGGTAATGC 51741
QY 482 AACCCGTTGCGATCTTCGGGCTCACCGAGTCTGTCAACGAGAGCGGGGGGGGATGC 541
DB 51742 GCCCGCGCGCGCTGTGTCTCAACGATCTTTCAGGCGGCGCGCTCCCGCGGAGA 51801
QY 542 CGGTGTCCGGGG 553
DB 51802 AGCAGCCCGCGG 51813

RESULT 4
US-10-329-148A-1/c
Sequence 1, Application US/10329148A
Publication No. US20040023343A1
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H
APPLICANT: Broughton, Mary C
APPLICANT: Crawford, Kathryn P
APPLICANT: Madhuri, Krishnamurthy
APPLICANT: Treadway, Patti J
APPLICANT: Turner, Jan R
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
FILE REFERENCE: 50489 D1/1
CURRENT APPLICATION NUMBER: US/10/329, 148A
PRIOR FILING DATE: 2002-12-23
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/603, 207B
PRIOR FILING DATE: EARLIER FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/370, 700
PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-09
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 80161
TYPE: DNA
ORGANISM: Saccharopolyspora spinosa
US-10-329-148A-1

Query Match 20.0%; Score 165.4; DB 17; Length 80161;
Best Local Similarity 52.8%; Pred. No. 1e-36;
Matches 386; Conservative 0; Mismatches 336; Indels 9; Gaps 1;

QY 15 CCGACCAATCGACGACAGGTTGGGAGATGTATGACTGTGTACCGCTTCTGAATCTC 74
DB 13575 CACACAGGATTTCCGACCGCGGATCAGGTGAGTCAATCTTGTGATGCTGTGGCGCACGG 13516
QY 75 GGTTCGGGGGGGCGCCCTTGGCGCATCAACAGCGCTACTATGGAGAAACAGCGCGGCTTC 134
DB 13515 GGTTCCTCTGACACACGTTTATCGGCGGCGGATTCGAGAGATGTCGTCGACACCC 13456
QY 135 CTGGACAGAGGCGGCGGACCGGCTCACCGACTTGTGCGGACGAGACCGGCTCGATG 194
DB 13455 GTGTGCGAGTGTGCGGACCACTGACGACTTGTTCATTCGACAGAGCGGCGCTCCGTCC 13396
QY 195 CCGGCTTCATCTGATGTGTGGGTGCGGTACCGGACACAGCAGCGCTGCGCGC 254
DB 13395 CCGAGCGCACCTGTTGACCTGGGCTGGCGGATGGGACGCGCTGATGCTGTGCGGCAATG 13336
QY 255 CCAACAACGATTCAGATCAACCGGATCAACGCTCAACGAGTGCAGTGGCCATGCGCGC 314
DB 13335 CCGCAGCGCGCTTTCAGATCAACGGAATCACGTTGAACGCCAGCATCTTCCGCGCCAC 13276
QY 315 TGAATTGCGAGCGGAACCGGACCTAAGCCACCGGCTGACTTCTGTCGTGATGTCAT 374
DB 13275 CAGGCTTCGCAACGAGACCGGACTGCGCGGCTGTGATGATCTGATGTCAGCGCGC 13216
QY 375 GTCCCTGCGGATCCCGGACATGCTTTGACGCGCGCTGGGCGCATGCTGCTTGTGA 434
DB 13215 CAGGCTGCGCTACCGGAGCGGTTCTTTCAGGCGCGGATGAGTCACTGCTGTCGA 13156
QY 435 GATGTCCGACCGGACCGGCTGATCCGGAATCTTTCAGATCACTCAACCGGCTGTCAT 494
DB 13155 GATGTGACACAGGCGCGCGGATTCGCGAGGTCAACCGAATCTTGAACCGGCGCGCG 13096
QY 495 CCTGGGCGTCAACGAGGTGTCAACAGAGGCGGCGGAGATCCGAGTGTCCGAGGA 554
DB 13095 GTTGTGCTTGGAGATATATCTGTTGATCTCCGGAAGTACCGCGGCTTTTG 13036
QY 555 CAGGTGCGGACCGGCTTTCGATGCTGCTGCTGAGCACTTTCGAAATGCTGCTGCTGC 614
DB 13035 GAGGCGGACAGACCGGCTATACCTT-----GAAACAGCTTCAACGCGCTGTCAGCA 12985
QY 615 AGCGGCTTTCGAGATCTTTCGATTTGGAGACGTTGTGTGAGGACCGGCTATCTTATGCC 674
DB 12984 AGCGGCTTTCGAGATCTTTCGAGTCAACGACTCACGAGCTCACGAGACGAGTGTCTTC 12925
QY 675 GCAGTTTCGCGGAGAGCTGCTGCGCACGAGCAGGATTCGCGGACGAGTACGAGCGCGG 734
DB 12924 CTGTACGTGACGAGATTCCTCGGAACTCGATGACTCTGCGGCTGAGGCTGCGGCG 12865
QY 735 TGTGCGCGGCT 745
DB 12864 TGTGCGCACT 12854

RESULT 5
US-10-844-716-1/c
Sequence 1, Application US/10844716
Publication No. US20050003409A1
GENERAL INFORMATION:
APPLICANT: Huang, Chengjin
APPLICANT: Chaleff, Deborah T.
APPLICANT: Ruppen, Mark E.
APPLICANT: Stephens, Jerome
TITLE OF INVENTION: Cloning Genes from Streptomyces Cyanogriseus Subsp.
FILE REFERENCE: AM100484
CURRENT APPLICATION NUMBER: US/10/844, 716
PRIOR FILING DATE: 2004-05-13
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn version 3.2

SEQ ID NO 1
LENGTH: 88400
TYPE: DNA
ORGANISM: bacteria
US-10-844-716-1

Query Match 17.0%; Score 141; DB 21; Length 88400;
Best Local Similarity 56.4%; Pred. No. 8.1e-30;
Matches 286; Conservative 0; Mismatches 215; Indels 6; Gaps 1;

QY 13 GGGCGACCAACATCCGACAGAGTTGGCGAGATGTATGACTGTGACCGCGTTGCTGAC 72
DB 51282 GGAATACCGACCCCGCGGACCTCGGTGCTATGACCGGATGACCGCTGTTGAC 51223
QY 73 TCGGTGCGCGGCGGCGCCCTGCGCCATCCACACGCGTACTGGAGAAACGACGGGCGGT 132
DB 51222 CGGGCCCTGGGCGAACAACGACCTCGGATGCTGGCCGACCGGACGACGACGAC 51163
QY 133 TCTTGACAGACGCGCGGACCGGCTGACCGGCTTGTGCGGAAACGACCGTGTGAT 192
DB 51162 CTGCGCCAGGCTCTCGGACCGCTGACCGGCTATGATGGGACGTTGCGGAAACAC 51103
QY 133 GG-----CGGGTTGACTGCTGATGTGGGTGCGGTACCGGACCAACGACGCTGCGC 246
DB 51102 GGGCGCCCGGTGACAGTGTGTGAGAGTGGCTGCGGCTCGGCGCGCGCTGCGCG 51043
QY 247 GTGCGCGGACCAACGCGGATCCAGATCACCGGATCACCGGTCAGGCGAGGTGCAATG 306
DB 51042 CTGGGCGACAGGACCGGTGACATCTGCGGATCACCATACGCCCCCGGAGGTGAG 50993
QY 307 ATGCGCGTGAATGCGGACGCGGACCGGCTGACCGGCTTGTGCGGTG 366
DB 50982 CTGGCGACCGCGCTGCGGACCGGTGCGGACTCGGAAACCGGCTTGTGAGTGGCGC 50923
QY 367 GATGCGATGTCCTGCGCGGACCGGACCAATGCTTTCAGCGCGGCTGGGCGCATGAGTGC 426
DB 50922 GACGCGATGACCTGCTCCCGACGCTCTTTCAGCGCGGTGCGGCGCTGAGTGC 50863
QY 427 CTGTTGAGATGTCGGAACCGGACCGGTGCAATCCGGGAAATCTTTCAGTACTCAAAACC 486
DB 50862 CTGCTGCAATGCTCCGACCGGCGCGGCTTTCAGGAGATGCGCGCTGTGCGCGCC 50803
QY 487 GGTGCGATCTCTGCGGCTGACCGGATC 513
DB 50802 GGTGCGCGCTGCGCGGCTGACCGGATC 50776

RESULT 6

US-10-156-761-925
Sequence 925, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OKURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 925
LENGTH: 849
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:

NAME/KEY: CDS
LOCATION: (1) .. (849)
US-10-156-761-925

Query Match 16.6%; Score 137.6; DB 15; Length 849;
Best Local Similarity 50.0%; Pred. No. 6.2e-29;
Matches 372; Conservative 0; Mismatches 369; Indels 3; Gaps 1;

QY 15 CGCACCAACATCCGACAGAGTTGGCGAGATGTATGACTGTGACCGCGTTGCTGAC 74
DB 24 CCGGCTCTTCTCTGAGTGGGAGATCTACGACCTGTGACCACTTATGATAC 83
QY 75 GGTGCGGCGGCGCCCTGCGCCATCCACACGCGTACTGGAGAACGACGGGCGGCTTC 134
DB 84 TGGCGTGGTGGAAACAACCACTCGATGCTGGCCGACGCGGAGGACGCGCAG---TTC 140
QY 135 CTGGCAGAGAGCGCGGCTGACCGGCTGACCGGCTTGTGCGGAAACGACCGTGTGATGG 194
DB 141 ACCGGGCAAGGCGCGGACCGGCTGACCGGCTTCTTCAATCGGCAAGCTGAGAGGCTATC 200
QY 195 CGGCGTTGACTGCTGATGTGGGATGCGGATCCGACCAACGCGCTGCGCGG 254
DB 201 GGGCGCGGAGTCTGAGACGTGCGGTGCGGATCCGGAAGCGCGGCTGCGGCTCCT 260
QY 255 CGAACACGCGATCCAGATCACCGGATCACCGGTCAGGCGAGGTGCAATGCGCGC 314
DB 261 GAGCGCGCGCTGATGTGTGCGGTGACGCGTGAAGGTGAGGTGCGGCTGCGGAC 320
QY 315 TGATTTGCGACGCGAAGCGGATCAAGCACCAGGTGATCTTCTGTGCTGTATGCTCAT 374
DB 321 CGCTCTGCGAAGAGTGCACGTTGCGGACCGGAGTGTGTTCAACCGTGCAGCGGAT 380
QY 375 GTCCCTGCGGATCCGACCAATGCTTTCAGACCGCGCTGCGGCAATGAGTGTGATGGA 434
DB 381 GGAATGCGCTTCTCCGACGAGTCTTTCAGACCGGCGGTGAGTGTCTCTGCA 440
QY 435 GATGTCGACACCGGACCGGTGCAATCCGGAAATCTTTCAGTACTCAAAACCGGTGCGAT 494
DB 441 CATGCCCGACCGCGCACAGGTGATCCGGAGATGCGCGGAGTGTCTCGCGCGCGGCGC 500
QY 495 CTTGCGGCTGACCGAGTGTGCAACGAGAACGCGGCGGCGGATGCGGCTGTGCGGGA 554
DB 501 GCTGCGCTGACGAGCTGCGACCTGCGGCGCTTTCGGGCGGACCGGACGAGAACCGGAG 560
QY 555 CAGGTGCGGACCGGCGCTTTCGATCTGCTGCGGCTGACCACTTTCGAAATGCTGCGTGC 614
DB 561 GTGCACTGCCAGTGTGCTGCGGCTGCGGCGCTGCTGACGAGTACGCGGAT 620
QY 615 AGCGGGGTGAGATCTCGATTTGGAGAGAGTGTCTGAGACCGCGTACTTATGCC 674
DB 621 GATGCGCGACCGCGGCTGGAATGCACTGAGCTGACCGACATCGGCGATCGAGTGTCTGCG 680
QY 675 GCAATTCGCGGAAGAGTGTGCGGCAACGACGAGGATCGGCGAGTACGGGCGGCG 734
DB 681 CCGCTTTCGCGCGGCTGCGGCTGCGGAGTACGAGTGAAGACGACCTGACGAGTACGCGGCGC 740
QY 735 TGTGCGGCGCTGCGGCGCGGCGGT 758
DB 741 CTTGCGGATCGGCGCTCGCGAGAT 764

RESULT 7

US-10-156-761-1/c
Sequence 1, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OKURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA


```

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

```

```

Query Match      16.6%; Score 137.6; DB 15; Length 9025608;
Best Local Similarity 50.0%; Pred. No. 8.8e-29;
Matches 372; Conservative 0; Mismatches 369; Indels 3; Gaps 1;

```

```

QY 15 CGACCAACATCGCAGAGTTGGGAGATGATGACCTTGACCGCGTTGCTGAATC 74
DB 1137457 CCGCCCTCTTCCCTGAGATGAGGAGCTAATGACCGTTTGAACGACTCAATGATCG 1137398
QY 75 GGTGCGGCGCGCGCCCTGCGCATCCACACGCGCTACTGGAGAGACGCGCGCGCTTC 134
DB 1137397 TGGCGTGGGTGGGAGAACCCACTCGGATATCTGGCGGACGCGGGGAGCGGAG---TTC 1137341
QY 135 CTGGCAGACGCGCGCGCGCGCTCACCGACCTTGTGCGCGAAGCGACCTGCTCGATGG 194
DB 1137340 ACCGGCAGAGCGCGCGACCGGCTCAACGATCTCTATCGGCAAGCTGAGAGGCATCAC 1137281
QY 195 CGGCGTTTCATCTGATGATGAGGCGGTGACCGGACACAGCGCTGCGCGTGGCGCG 254
DB 1137280 GGGCGCGCGGCTCTGACCTGCGCTGCGGTTCCGGAAGCGCGGCTGGCGCTCGCCT 1137221
QY 255 CGACCAACGCGATCCAGATCAACCGGCATCACCGTCAGCGACGAGTGAAGTGCATCGCGC 314
DB 1137220 GAGCGCGCGCGGTGATGATCTCGCGCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1137161
QY 315 TGATGGGACGCGGACGCGGACCTAAGCACCGGCTGAGCTTCTGCTGCTGATGCCAT 374
DB 1137160 CGCTCTCGCAAGACAGTCGACGCTGCGGACCGGCTGCTGCTGCTGCTGCTGCTGCA 1137101
QY 375 GTCCCTGCGCTACCGGACCAATGCTTTCGACGCGCGCTGGCCATGAGTGCCTGTTGGA 434
DB 1137100 GGAACCTGCGCTTTCGCGGCTCTTTCGACGCGCGCTGGCCCTGAGTGTCTCTGCA 1137041
QY 435 GATGTCGGAACCGGACCGGTCCATCCGGAATCTCTTCAGATCTCAAAACCGGTGCGAT 494
DB 1137040 CATGCCAAGCTCCGACAGTGAATCCGGAGATGCCCCGGGTGTCCGCCCGCGCGCG 1136981
QY 495 CCTGCGGTCACCGGAGTCTCAAAAGAGAGCGCGCGCGGATGCCGCTGTCGCGGGA 554
DB 1136980 GCTGCGCGTCACGAGCTGCGACCTGCGCGCTTTCGCGGACCGGCAATGAAAGCGCGGGA 1136921
QY 555 CAGGTGCGACCGGCTTTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 614
DB 1136920 GTGACGCTCCCACTGCTGCGGCTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1136861
QY 615 AGCGGAGTTCAGATCTCTGATGAGGAGAGCTGCTGCTGAGAGACCGCGTACTTTCATGCC 674
DB 1136860 GATGCGCGAGCGCGCGGCTGGAACTGATGAGCTGACCGACATGCGGATCAGAGTCTGCGG 1136801
QY 675 GCACTTTCGCGAGAGCTCGCTGCGACACGACGAGGATGCGGACAGGTAACGCGCGCGC 734
DB 1136800 CCCCCTCTTTCGCGCGCGCTGCGTGAACAGTGAACGACACCTCGACGATGACGCGCGCG 1136741
QY 735 TGTGCGCGGCTGCGCGCGCGCGGT 758

```

```

DB 1136740 CTTGCGGATCGGCGTCCGCGAGAT 1136717

```

```

RESULT 8
US-09-953-348-36

```

```

; Sequence 36, Application US/09953348
; Publication No. US20030134398A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, David. H
; APPLICANT: Mao, Yingqiang
; APPLICANT: Varoglu, Mustafa
; APPLICANT: He, Min
; APPLICANT: Sheldon, Paul

```

```

; TITLE OF INVENTION: MITOMYCIN BIOSYNTHETIC GENE CLUSTER

```

```

; FILE REFERENCE: 600.530US1
; CURRENT APPLICATION NUMBER: US/09/953,348
; CURRENT FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: PCT/US00/06394
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/266965
; PRIOR FILING DATE: 1999-03-12
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 153
; SEQ ID NO 36
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Streptomyces lavendulae
US-09-953-348-36

```

```

Query Match      16.1%; Score 133.2; DB 10; Length 852;
Best Local Similarity 54.0%; Pred. No. 1.1e-27;
Matches 295; Conservative 0; Mismatches 248; Indels 3; Gaps 1;

```

```

QY 15 CGACCAACATCGCAGAGTTGGGAGATGATGACCTTGACCGCGTTGCTGAATC 74
DB 36 CTCACCGCGCTCGAGAGAGTGGGCGGCTTAACGACCGGTTCAACGCGCTGGAGCCGC 95
QY 75 GGTGCGGCGCGCGCCCTGCGCATCCACACGCGCTACTGGAGAGAGAGAGAGAGAGAGAGAG 134
DB 96 CTCCCTGCGCGAGAACCTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 152
QY 135 CTGGCAGAGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 194
DB 153 GCTGCGCGAGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 212
QY 195 CGGCGTTTCATCTGATGATGAGGCGGTGACCGGACCAAGCGCTGCGCGCTGCGCG 254
DB 213 CGGCTCGCGGCTCTGACCTCGGCTGCGGCGGCGGAGCCCGGCGTACGATCGCGCG 272
QY 255 CGACCAACGCGATCCAGATCAACCGGCATCACCGTCAGCGAGTGAAGTGAAGTGAAGTGAAG 314
DB 273 GCTCAGCGGAGCGCATGTCACGGGATCTGCGTGAAGCCATGAGCAAGTGTCTCGGAGCAA 332
QY 315 TGATGGGACGCGGACGCGGACCTAAGCACCGGCTGAGCTTCTGCTGCTGCTGCTGCTGCTG 374
DB 333 CGGCGTGGCGGAGAGCGCGGCTCGCCGACCGGCGCGCTTTCGAGCGGCGCGAGCGAT 392
QY 375 GTCCCTGCGTACCGGACCAATGCTTTCGACGCGCGCTGAGGCGGCTGAGTGTCTGTTGGA 434
DB 393 GACCTCCCTTGAAGAGAGAGAGCTTTCGACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 452
QY 435 GATGTCGGAACCGGACCGGTCCATCCGGAATCTCTTCAGATCTCAAAACCGGTGCGAT 494
DB 453 CATGCCAAGCTCCGACAGTGAATCCGGAGATGCCCCGGGTGTCCGCCCGCGCGCGAG 512
QY 495 CCTGCGGTCACCGGAGTCTCAAAAGAGAGCGCGCGGATGCCGATGTCGCGGGA 554
DB 513 TCTGAGCTACCGACTTCTTGAAGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 572
QY 555 CAGGTG 560
DB 573 CAGCG 578

```

RESULT 9
US-10-267-255-36
; Sequence 36, Application US/10267255
; Publication No. US20030124689A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D
; APPLICANT: Mao, Y
; APPLICANT: Varoglu, M
; APPLICANT: He, M
; APPLICANT: Sheldon, P
; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
; FILE REFERENCE: 600,456US1
; CURRENT APPLICATION NUMBER: US/10/267,255
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 09/266,965
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 08/624,447
; PRIOR FILING DATE: 1996-08-19
; PRIOR APPLICATION NUMBER: PCT/US94/11279
; PRIOR FILING DATE: 1994-10-06
; PRIOR APPLICATION NUMBER: US 08/133,963
; PRIOR FILING DATE: 1993-10-07
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 36
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Streptomyces lavendulae
US-10-267-255-36

Query Match 16.1%; Score 133.2; DB 15; Length 852;
Best Local Similarity 54.0%; Pred. No. 1.1e-27;
Matches 295; Conservative 0; Mismatches 248; Indels 3; Gaps 1;
QY 15 CGCACCAATCGCAGCGAGTTGGCAGATGATGATCTGCTGCAAGCCGTTGCTGAATC 74
DB 36 CTCACCCGCTCCGAGAGTGGCGCGCTCTACACACCGGTTCAACCGGCTGGAGCCGC 95
QY 75 GGTGCGGCGGCGCCCTGCGGCATCACACCGGCTACTGGAGAGACGAGCGGCGGCTTC 134
DB 96 CTCCTCGGCGAGAACCTGCACTTGGCTACTGGAACT---CCCCGACAGCAGATGCC 152
QY 135 CTGGCAGACGAGCGCGCGACCGGCTGACCGACTTGTGCGCGACGAGCCGTGCTGATG 194
DB 153 GCTGACCGAGGCGACCGACCGGCTCACCGACATGATGGCCGAGCGGCTGCCGATCG 212
QY 195 CGGCGTTGATGCTGCTGATGAGGGGTGGGATACCGGACAACTCAGCGCTGCGGCGCG 254
DB 213 CGGCTCCGCGCTCTGACCTCGCGTGGCGCGTGGGAGACCCCGGCGATGCGATGCCG 272
QY 255 CGACACGCGATCCAGATCACCGGCATCACCGCTCAGCGAGGACAGTGGCCATGCGCGC 314
DB 273 GCTCAGCGAGGCGATGTCACGGGCACTCGGTAGCCATGAGAGGTCTGTCGGGCGCA 332
QY 315 TGATTGCGACGCGACCGGACCTAAGCCACCGGGTGAATTCTGTGCGGTGATGCCAT 374
DB 333 CGCGCTGACGAGGAGCGCGGCTCGCGACCGGGCGCGCTTCAGCGGGCGAGCGCAT 392
QY 375 GTCCCTGCGATACCGGACATGCTTTCAGCGCGCGCTGGGCGCATGCGTGGCTTGA 434
DB 393 GGAACCTCCCTTCAGGAGCGAGAGCTTCAGCGCGCTTCGCGCTCGATCGATCATCA 452
QY 435 GATGTCGGAACCGGACCGGTCATCCGGGAAATCTTGAGATCTCAAAACCGGTGCGAT 494
DB 453 CATGCCGACCGCGCGCAGGTGCTGCCAGGTGCGCGGGTCTGCGGCGCGGAGGCGCG 512
QY 495 CTTGCGGCTCACCGAGGTGCTCAACGAGAACCGGCGCGGAGTCCGATGCTCGGAGGA 554
DB 513 TCTGTGCTCACCGACTTCTTCGAGCGGCGCGCGCGCGCGCGCGCGCGCGT 572
QY 555 CAGGTG 560
| | | | |

DB 573 CCAGCG 578
RESULT 10
US-09-953-348-76/c
; Sequence 76, Application US/09953348
; Publication No. US20030134398A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, David, H
; APPLICANT: Mao, Yingying
; APPLICANT: Varoglu, Mustafa
; APPLICANT: He, Min
; APPLICANT: Sheldon, Paul
; TITLE OF INVENTION: MITOMYCIN BIOSYNTHETIC GENE CLUSTER
; FILE REFERENCE: 600,530US1
; CURRENT APPLICATION NUMBER: US/09/953,348
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: PCT/US00/06394
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/266965
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 76
; LENGTH: 53500
; TYPE: DNA
; ORGANISM: Streptomyces lavendulae
US-09-953-348-76

Query Match 16.1%; Score 133.2; DB 10; Length 53500;
Best Local Similarity 54.0%; Pred. No. 1.3e-27;
Matches 295; Conservative 0; Mismatches 248; Indels 3; Gaps 1;
QY 15 CGCACCAATCGCAGCGAGTTGGCAGATGATGATCTGCTGCAAGCCGTTGCTGAATC 74
DB 7549 CTCACCCGCTCCGAGAGTGGCGCGCTCTACACCGGTTCAACCGGCTGGAGCCGC 7490
QY 75 GGTGCGGCGGCGCCCTGCGGCATCACACCGGCTACTGGAGAGACGAGCGGCGGCTTC 134
DB 7489 CTCCTCGGCGAGAACCTGCACTTGGCTACTGGAACT---CCCCGACAGCAGATGCC 7433
QY 135 CTGGCAGACGAGCGCGCGACCGGCTGACCGACTTGTGCGCGACGAGCCGTGCTGATG 194
DB 7432 GCTGACCGAGGCGACCGACCGGCTCACCGACATGATGGCGAGCGGCTGCGCATCG 7373
QY 195 CGGCGTTGATGCTGCTGATGAGGGGTGGGATACCGGACAACTCAGCGCTGCGGCGCG 254
DB 7372 CGGCTCCGCGTCTGAGCTCTGCTGCGGTGGGAGACCCCGGCGTACGATGCCGCG 7313
QY 255 CGACACGCGATCCAGATCACCGGCATCACCGCTCAGCGAGGTGACGATGCGCGC 314
DB 7312 GCTCAGCGAGCGCATGTCACGGGCACTCGGTAGCCATGAGAGGTCTGTCGGGCGCA 7253
QY 315 TGATTGCGACGCGACCGGACCTAAGCCACCGGCTGAGACTTCTGCTGCGTGCATGCCAT 374
DB 7252 CGCGCTGACGAGGAGCGCGGCTCGCGACCGGGCGCGCTTCAGCGGGCGAGCGCAT 7193
QY 375 GTCCCTGCGATACCGGACCAATGCTTTCAGCGCGCGCTGGGCGCATGCGTGGTGA 434
DB 7192 GGAACCTCCCTTCAGGAGCGAGACTTCAGCGCGCTTCATGCGCTCGAATGATCATCA 7133
QY 435 GATGTCGGAACCGGACCGGTGCTCCGGGAAATCTTCGATCTCAAAACCGGTGCGAT 494
DB 7132 CATGCCGACCGCGCGCAGGTGCTGCCAGGTGCGCGGGTGTGCGGCGCGGAGGCGCG 7073
QY 495 CTTGCGGCTCACCGAGGTGCTCAACGAGAACCGGCGGCGGAGTCCGATGCTCGGAGGA 554
DB 7072 TCTGTGCTCACCGACTTCTTCGAGCGGCGCGCGCGCGCGCGCGCGCGCGT 7013
QY 555 CAGGTG 560
DB 7012 CCAGCG 7007
| | | | |

RESULT 11
US-10-267-255-76/c
; Sequence 76, Application US/10267255
; Publication No. US20030124689A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D
; APPLICANT: Mao, Y
; APPLICANT: Varoglu, M
; APPLICANT: He, M
; APPLICANT: Sheldon, P
; TITLE OF INVENTION: Mitomycin bioisynthetic gene cluster
; FILE REFERENCE: 600 456US1
; CURRENT APPLICATION NUMBER: US/10/267,255
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 09/266,965
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 08/624,447
; PRIOR FILING DATE: 1996-08-19
; PRIOR APPLICATION NUMBER: PCT/US94/11279
; PRIOR FILING DATE: 1994-10-06
; PRIOR APPLICATION NUMBER: US 08/133,963
; PRIOR FILING DATE: 1993-10-07
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 76
; LENGTH: 53500
; TYPE: DNA
; ORGANISM: Streptomyces lavendulae
US-10-267-255-76

Query Match 16.1%; Score 133.2; DB 15; Length 53500;
Best Local Similarity 54.0%; Pred. No. 1,3e-27;
Matches 295; Conservative 0; Mismatches 248; Indels 3; Gaps 1;

QY 15 CGCAACAATCGCAGAGTGGGCAAGATGATGACCTGTACGCGCTGTGCTGAATC 74
DB 7549 CTACCCGCTCGAGAGATGGGCGGCTCTACGACCGGTTACCGCGTGGAGACCG 7490
QY 75 GGTGCGGGGCGGCGCTGCGCATTCACACGCGCTACGGAGAGACGAGGCGGCTTC 134
DB 7489 CTCCTCGGGGAGAACCTGCACTTCGGCTACTGGAACT---CCCCGACAGCAGGAGGCG 7433
QY 135 CTGGCAGACGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 194
DB 7432 GCTGGCGAGGCGCAGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7373
QY 195 CGGCGTTGCACTGCTGATGAGGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 254
DB 7372 CGGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7313
QY 255 CGCAACAATCGCAGAGTGGGCAAGATGATGACCTGTACGCGCTGTGCTGAATC 314
DB 7312 GCTCAGCGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7253
QY 315 TGATGCGACGCGGAGCGGAGCTTAACCAACCGGAGTGTCTGCTGCTGCTGCTGCT 374
DB 7252 CGGCTGCGAGGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7193
QY 375 GTCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 434
DB 7192 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7133
QY 435 GATGCTGCGACGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 494
DB 7132 CATGCGCGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7073
QY 495 CTTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 554
DB 7072 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7013
QY 555 CAGGTG 560

Db 7012 CCAGCG 7007
RESULT 12
US-09-980-217-1/c
; Sequence 1, Application US/09980217
; Publication No. US20040219645A1
; GENERAL INFORMATION:
; APPLICANT: Biotech Technology Limited
; APPLICANT: Leadlay, Peter F
; APPLICANT: Staunton, James
; APPLICANT: Olynyk, Marko
; TITLE OF INVENTION: Polyketides and their synthesis
; FILE REFERENCE: IS/BP5858469
; CURRENT APPLICATION NUMBER: US/09/980,217
; PRIOR FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: PCT/GB00/02072
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: GB 9912563.5
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 30000
; TYPE: DNA
; ORGANISM: Streptomyces cinamomensis
; FEATURE:
; OTHER INFORMATION: Nucleotides 1 to 30000 of the monensin
US-09-980-217-1

Query Match 14.5%; Score 120.2; DB 11; Length 30000;
Best Local Similarity 52.5%; Pred. No. 5.8e-24;
Matches 263; Conservative 0; Mismatches 238; Indels 0; Gaps 0;

QY 50 ACCTGCTGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 109
DB 10374 ACAAGCTTTCAGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 10315
QY 110 ACTGGAGAGACGAGCGGCGGCTTCTGCGACAGAGCGGCGGCTGCTGCTGCTGCT 169
DB 10314 TCGACGCGCGGCGAGCAGCGGCGGCTTCTGCGACAGAGCGGCGGCTGCTGCTG 10255
QY 170 TCGCGGAAGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 229
DB 10254 TGATCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 10195
QY 230 GACCAACAAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 289
DB 10194 GCAACCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 10135
QY 290 GCCAGGTGAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 349
DB 10134 GCGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 10075
QY 350 TGAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 409
DB 10074 TCGCTTTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 10015
QY 410 CTTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 469
DB 10014 GCTGCGGCTTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9955
QY 470 TTGAGTACTCAAAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 529
DB 9954 ACCGCTGCTGAGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9895
QY 530 GCGCGGAGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 550
DB 9894 ATCCAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9874

RESULT 13

Query Match	11.3%	Score 93.8	DB 19	Length 1773
Best Local Similarity	52.7%	Pred. No. 1.5e-16		
Matches 203; Conservative	0	Mismatches 182	Indels 0	Gaps 0

Qy	133	TCCTGGACAGACAGCGCCGACCCGGCTCAACCACTTGTGTCCGGAACGACCCGTGCTGCAT	192
Db	1268	TTCCACCGCGAGGCCACGCGCGTTCACAGAGAGCCGCTGCGACCTTCTTCAAGGCCAAG	1208
Qy	193	GCGCGCGTTTCAGTACTCTCGATGTGGGGTGTGGGTATCCGGACAACAGCCCTGGCGCTCGCG	252
Db	1208	CCGGGACACCCGCTCTCGACGTGGGCTGGCGGCGCTGCGCGGGCCCATGCGGCCATTCGCG	1144
Qy	253	CGCGACAACGCGGATCCAGATCAACCGGCATCAACGTTCAGCCAGGTGCAAAGTGCATGTGCC	312
Db	1148	GCCCACTCCGGGCTCCAGCTCGTGGGCAATCAACAGATACCAAGTACCAAGTGAACCGCGCC	1088
Qy	313	GCTGATTTGGCAACGGGAAACGGGACTAAGCCACCGGGTGGGACTTCTCTGTGCTCGATGCGC	372
Db	1088	CGCGGCAACAACCGGAAAGCGGGCTTCGACTCCCGGCTGCGAGGTGTCTGGCGTAACTTTC	1028
Qy	373	ATGTCCCTACGCGTACCCGGACCAATGTTTTCGACGCGCGCTGGGGCATGCAAGTGCCTGTGG	432
Db	1028	CTCTTCATGCGGTTCTCCGACGCGCTCTTTCGACGGCGGCTACTCATGAGGCCACCTGCG	969
Qy	433	GAGATGTCGAAACCGGACCGTGCATTCGGGAAATCTTTCGAGTACTCAAAACCCGGTGGC	492
Db	968	CACGGGCGCAAGCTGACAGACGTCTAACGCGGAGGTCTTCCGGGTGCTCAAGCCGGGCGGC	909
Qy	493	ATCTCTGGCGTACCGGAGTGTGTCA	517
Db	908	CTTCAAGTCTCTTACGAGTGGGTCA	884

US-10-259-194A-251

Sequence 251, Application US/10259194A

Publication No. US20040010815A1

GENERAL INFORMATION:

APPLICANT: Lange, Markus B.

APPLICANT: Gassemeian, Majid

APPLICANT: Briggs, Steven P.

APPLICANT: Cooper, Bret

APPLICANT: Glazebrook, Jane

APPLICANT: Golf, Stephen A.

APPLICANT: Katagiri, Fumiyuki

APPLICANT: Krepis, Joel

APPLICANT: Moughamer, Todd

Query Match	11.3%	Score 93.2	DB 17	Length 1092
Best Local Similarity	52.6%	Pred. No. 2.2e-16		
Matches 203; Conservative	0	Mismatches 183	Indels 0	Gaps 0

Qy	13	TTCTGTGGCAGACAGGCGCGCGGACCGGCTCAGCACTTGTGCGCGCAACGGAACCGGTGCA	19.1
Dp	321	TTTCCACCCGGAGGCCACCGCGCTTCACGAGAGGCGGTGCGCGCACTCTTCACAGGCCAA	380
Qy	192	TGGCGGCGATTTCGACTGCTGATGTGGGGGTGGGTTACCGGACCAACCAACGCTGCGCGTGC	251
Dp	381	GCGGGGACACCGGCTCTTGACGTGTGGGCTGGGGGGTGGGGGGCCAAATGGCGGCATCGC	440
Qy	252	GCGCGACAAACGGGATTCAGATCACCGGATCACCGTCAAGCCAGGTGCAATGGCCATCGC	311
Dp	441	CGCACCTCTCGGCTCCACAGTGTGGGATCACCATCAACAGTAAACAGGTGAAACCGCGC	500
Qy	312	CGGTGATTTGGCAGACGGGAAACCGGACCTPAAGCAACCGGGTGGACTTTCGTGGCGTCGATGC	371
Dp	501	CCGCGCGACAAACCGGAGAGCGGGCTGTGACTCCCGCTGCGAGGTGCTTGTGGCGTAACTTT	560
Qy	372	CATGTCCCTGCGGTTACCCGGACAAATGCTTTGACGCGCGCTGGGCGCAATGACGTGCGTGT	431
Dp	561	CCCTCTCAATGCGGTTCTCCGAGGCTCTCTTTCGACGGCGGCTCACTCACTCGAGGCGCACTG	620
Qy	432	GGAGATGTCGGAACCGGACCGTCCATCCGGGAAATCTTCGATGACTCAAAACCCGCTGG	491
Dp	621	CCAGCGGCCCAAGCTGACAGAACGTCTACGGCGAGGCTTCCGGCTGCTCAAGCCGGGCGG	680
Qy	492	CATCTCTGGGCTTACCGGAGGTGCTGA	517
Dp	681	CCCTTAACGTCTCTCAAGATGGGTTCA	706

RESULT 15
US-10-259-194A-529
Sequence 529, Application US/10259194A
Publication No. US20040010815A1
GENERAL INFORMATION:
APPLICANT: Lange, Markus B.
APPLICANT: Chasseman, Majid
APPLICANT: Briggs, Steven P.
APPLICANT: Cooper, Bret
APPLICANT: Glazebrook, Jane
APPLICANT: Golf, Stephen A.
APPLICANT: Katagiri, Fumiyuki
APPLICANT: Kreps, Joel
APPLICANT: Moughamer, Todd
APPLICANT: Provart, Nicholas
APPLICANT: Ricke, Darrell
APPLICANT: Zhu, Tong
TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF PLANT GENES
FILE REFERENCE: 70029-NP

```

:
: CURRENT APPLICATION NUMBER: US/10/259,194A
:
: CURRENT FILING DATE: 2003-01-07
:
: PRIOR APPLICATION NUMBER: US 60/325,277
:
: PRIOR FILING DATE: 2001-09-26
:
: PRIOR APPLICATION NUMBER: US 60/370,743
:
: PRIOR FILING DATE: 2002-04-04
:
: PRIOR APPLICATION NUMBER: US 60/370,620
:
: PRIOR FILING DATE: 2002-04-04
:
: NUMBER OF SEQ ID NOS: 662
:
: SOFTWARE: Patentlist.pl version 3.0.4 (C) 2001 Syngenta
:
: SEQ ID NO 529
:
: LENGTH: 1418
:
: TYPE: DNA
:
: ORGANISM: Trifolium aestivum
:
: US-10-259-194A-529

```

Query Match	10.8%;	Score 89.4;	DB 17;	Length 1418;
Best Local Similarity	52.2%;	Pred. No. 2.6e-15;		
Matches 198; Conservative	0;	Mismatches 181;	Indels 0;	Gaps 0;

OY 139 TAGCAGGGCCGCGCACCGGCTCACCGACTTGTGCGCGGAACGGACCGGTGCTCGATGCGGGC 198
 Db 408 CCGGAGGCGCACCGCGGGTGCACGAGGAGCGCGTGGCCGACTCTGCTCGAAGGCGCGCCCCGGG 467
 OY 199 GTTCGACTGCTCGATGTGGGATGTGGGTCACGGACACCAAGCGCTGCGCGCGTGGCGCGGAC 258
 Db 468 AAGCGCTGTGTCGAGCGTGGCTGGCGCGCTGCGGGGCCCATGTGCGGCCCATGCGGCGCAC 527
 OY 259 AAGCGGATTCAGATCACCGGCATCACCGTTCAGCGCAAGTGCATAGTGGCCATCGCCGCTGAT 318
 Db 528 TCGGCTCTCGAAGTGTGTCGGGATTCACATCAACAGATCACAGGTTCACCGCGCCGCTCG 587
 OY 319 TGGCAGCGGAAACGGGCACTAAGCCACCGGGTGGATCTTCGAGCGGTGCATGCAATGCC 378
 Db 588 CACAAACCGGAAGCGCGGCTTCGACGCGCAATGCCAGAGTGTGTGCGGCAACTTCATATGCC 647
 OY 379 CTGCGGTAACCCGAGACATCTTTCGACGCGCGCTGGGCCATGCAAGTGTGCTGTGGAGATG 438
 Db 648 ATGCCGTTCCGGAGCGCCTCTTCGACGCGGCGCTTACCTCATCGAAGGCACCTGCGACGCG 707
 OY 439 TTCGAAACGGAGCCGTGCCATCCGGGAAATCTTCGAGTACTCAAAACCGGTGGCATCTTC 498
 Db 708 CCGAGGCTGCAGAGACCTTCACGGCGGAGGTGTTCCGCGTGTCTCAAGCCAGGGGGCTCTTAC 767
 OY 499 GGGCTCACAGAGTGTCTCA 517
 Db 768 GTCTCTTACGAGTGGGTCA 786

Search completed: August 29, 2005, 17:23:43
Job time : 718 secs

This Page Blank (uspto)

OY		61	CGGTGCTGAAC	TCCGTCGGCGGGCGGCCCTCCTGCATTCACCAAGGCTA	CTGGAGAAG	120
Db		61	CGGTGCTGAAC	TCCGTCGGCGGGCGGCCCTCCTGCATTCACCAAGGCTA	CTGGAGAAG	120
OY		121	GACGGGCGGGCTTCTCC	TGGCAGAGCGCGCGCATCACCGACTTCCTGCGCGAACG		180
Db		121	GACGGGCGGGCTTCTCC	TGGCAGAGCGCGCGCATCACCGACTTCCTGCGCGAACG		180
OY		181	ACCCTGCTGATG	GGCGGCTTTGACTGCTGATG	TG	240
Db		181	ACCCTGCTGATG	GGCGGCTTTGACTGCTGATG	TG	240
OY		241	CTGCGCGTGCGCGCGACAACCGCATTCAGATCA	CCGCGATCACCGGATCACCGCAGGTCAA		300
Db		241	CTGCGCGTGCGCGCGACAACCGCATTCAGATCA	CCGCGATCACCGGATCACCGCAGGTCAA		300
OY		301	GTGGCCATCGCGCGTATTGCGCACCGGACCGGAACTAAGCCAC	CGGGTGGACTTCTCG		360
Db		301	GTGGCCATCGCGCGTATTGCGCACCGGACCGGAACTAAGCCAC	CGGGTGGACTTCTCG		360
OY		361	TGGCGTCAGTCCATGCTCCCGCGATACCGCGA	CATA	TGCTTTGAGCGCGCTTGGCCATG	420
Db		361	TGGCGTCAGTCCATGCTCCCGCGATACCGCGA	CATA	TGCTTTGAGCGCGCTTGGCCATG	420
OY		421	CAGTGCCTGTGAGATG	TCCGACCGGACCGGTGCATTCGGGAAATCCTT	GAGTAATC	480
Db		421	CAGTGCCTGTGAGATG	TCCGACCGGACCGGTGCATTCGGGAAATCCTT	GAGTAATC	480
OY		481	AAAACCCGGTGCGATCTCTGGCGCTACCGAGGTCGTCAA	CGGAAGCGGGCGCGGGATG		540
Db		481	AAAACCCGGTGCGATCTCTGGCGCTACCGAGGTCGTCAA	CGGAAGCGGGCGCGGGATG		540
OY		541	CCGGTGTCCGGAGAGGTGCGCGACCGGCTTTCGATCTGCGTGGTAGCA	ATTCTG		600
Db		541	CCGGTGTCCGGAGAGGTGCGCGACCGGCTTTCGATCTGCGTGGTAGCA	ATTCTG		600
OY		601	GAAATGCTGCGTGCAGACCGGGTTGCA	GATCCTGATTTGGAGAGACGTGTCTGCGAGACC		660
Db		601	GAAATGCTGCGTGCAGACCGGGTTGCA	GATCCTGATTTGGAGAGACGTGTCTGCGAGACC		660
OY		661	CGGTACTTCATGCGCAGTGTGCGCGAAGAGCTGTGCGGAC	CACACAGGGATGCGCGAC		720
Db		661	CGGTACTTCATGCGCAGTGTGCGCGAAGAGCTGTGCGGAC	CACACAGGGATGCGCGAC		720
OY		721	AGGTACGGGCGGCTGTGCGCGGCTGGGCGCGCGGTCTGGATTATGAGAAATATCC			780
Db		721	AGGTACGGGCGGCTGTGCGCGGCTGGGCGCGCGGTCTGGATTATGAGAAATATCC			780
OY		781	CACGACATGCGGTATGCAATTCAGACGGCGCGAAGCCGGTGGCTGA			828
Db		781	CACGACATGCGGTATGCAATTCAGACGGCGCGAAGCCGGTGGCTGA			828
RESULT 2						
AX089419/c						
LOCUS	AX089419		45624 bp	DNA	linear	PAT 21-MAR-2001
DEFINITION	Sequence 4 from Patent WO0116303.					
VERSION	AX089419					
KEYWORDS	AX089419.1 GI:13443680					
SOURCE	.					
ORGANISM	Saccharopolyspora spinosa Saccharopolyspora spinosa Bacteria; Actinobacterii; Actinobacteridae; Actinomycetales; Pseudonocardineae; Pseudonocardiales; Saccharopolyspora.					
REFERENCE	Eberz, G., Moehrle, V., Froede, R., Velten, R. and Salas, J.A. Nucleic acids which code for the enzyme activities of the spinosyn biosynthesis Patent: WO 0116303-A 4 08-MAR-2001;					
AUTHORS	JOURNAL					
TITLE	BAYER AG (DE)					
FEATURES	LOCATION/Qualifiers					
SOURCE	1..45624					

```

/organism="Saccharopolyspora spinosa"
/mol_type="unassigned DNA"
/db_xref="taxon:60894"

```

Query Match	100.0%	Score 828	DB 6	Length 45624
Best Local Similarity	100.0%	Pred. NC 0		
Matches 828, Conservative	0	Mismatches 0	Indels 0	Gaps 0

QY	1	GTGTTGCCAGGATGGCGCAACCAACATCCGAGAGAGTTGGGCGAGATGTAAGACTGTGCACG	60
Db	828	GTGTTGCCAGGATGGCGCAACCAACATCCGAGAGAGTTGGGCGAGATGTAAGACTGTGCACG	769
QY	61	CCGTTGTGAACTCGGTGCGCGGCGGCCCTCTGCGCATTCACACACGCGTACTGGGAGAAC	120
Db	768	CCGTTGTGAACTCGGTGCGCGGCGGCCCTCTGCGCATTCACACACGCGTACTGGGAGAAC	709
QY	121	GACGGGCGGGGCTTCTCTGGCAGCAGGCGCGCCGACCGGCTCACCAGACCTTTGTGCCGAGCG	180
Db	708	GACGGGCGGGGCTTCTCTGGCAGCAGGCGCGCCGACCGGCTCACCAGACCTTTGTGCCGAGCG	649
QY	181	ACCNGCTCGAATGGGCGGGGTTTCGACTGTGATGTGGGGTGGGATCCGACCAACCAAGCG	240
Db	648	ACCNGCTCGAATGGGCGGGGTTTCGACTGTGATGTGGGGTGGGATCCGACCAACCAAGCG	589
QY	241	CTGCGCGGTGCGCGCGACCAACGCGCATCCAGATCACCGGCATCACCGTTCAGCCAGATGCA	300
Db	588	CTGCGCGGTGCGCGCGCGACCAACGCGCATCCAGATCACCGGCATCACCGTTCAGCCAGATGCA	529
QY	301	GTGGCCATTCGGCGCGTGATTTGGGCACCGGAAACGGCGAATTAAGCCACCGGGTGGACTTCTCG	360
Db	528	GTGGCCATTCGGCGCGTGATTTGGGCACCGGAAACGGCGAATTAAGCCACCGGGTGGACTTCTCG	469
QY	361	TGCGTGCATGTCATGTCCCTGCGCTGACCCGACATGCTTTTCGACGCGCGCTGGGCGCATG	420
Db	468	TGCGTGCATGTCATGTCCCTGCGCTGACCCGACATGCTTTTCGACGCGCGCTGGGCGCATG	409
QY	421	CAGTGCCTGTTTGGAGATGTCTCCGAACCCGACCCGTGTCCATCCGGGAAATCTTTTCGAGTACTC	480
Db	408	CAGTGCCTGTTTGGAGATGTCTCCGAACCCGACCCGTGTCCATCCGGGAAATCTTTTCGAGTACTC	349
QY	481	AAACCCCGGTGGCATCTCTCGCGGTACACCGAGGTGTCGCAACAGAGAAGCGGGCGGGGATG	540
Db	348	AAACCCCGGTGGCATCTCTCGCGGTACACCGAGGTGTCGCAACAGAGAAGCGGGCGGGGATG	289
QY	541	CCGGTGTCCGGGGACAGGTGGCCGACCGGCGCTTTCGATCTTGCCTGCTGAGCAACTTCTG	600
Db	288	CCGGTGTCCGGGGACAGGTGGCCGACCGGCGCTTTCGATCTTGCCTGCTGAGCAACTTCTG	229
QY	601	GAATTCGCTGCGTGTGAGGGGGGTTTGAAGATCTCTGATTTGGAGAGACGTGTCTGAGAACCC	660
Db	228	GAATTCGCTGCGTGTGAGGGGGGTTTGAAGATCTCTGATTTGGAGAGACGTGTCTGAGAACCC	169
QY	661	CGGACTTTCATGCGCGCAAGTTGCGCGAAGAGCTCGCTCGCACACGACAGGAGATGCGCGAC	720
Db	168	CGGACTTTCATGCGCGCAAGTTGCGCGAAGAGCTCGCTCGCACACGACAGGAGATGCGCGAC	109
QY	721	AGGTACCGGCGCGCTGTGCGCGGCTGGGACCGCGCGTCTTGCGATTATGAGAAATATGCC	780
Db	108	AGGTACCGGCGCGCTGTGCGCGGCTGGGACCGCGCGTCTTGCGATTATGAGAAATATGCC	49
QY	781	CACGACATGGGCTATGTGGAATTCTGACGGGCGCGAAGCGGTCGGGCTGA 828	
Db	48	CACGACATGGGCTATGTGGAATTCTGACGGGCGCGAAGCGGTCGGGCTGA 1	
RESULT 3			
LOCUS	AX089416	50000 bp	DNA linear PAT 21-MAR-2001
DEFINITION	Sequence 1 from Patent WO0116303.		
ACCESSION	AX089416		
VERSION	AX089416.1	GI:13443677	
KEYWORDS	.		

SOURCE Saccharopolyspora spinosa
 ORGANISM Saccharopolyspora spinosa
 Bacteria; Actinobacteridae; Actinomycetales;
 Pseudonocardineae; Pseudonocardaceae; Saccharopolyspora.

REFERENCE 1
 Eberz G., Moenle V., Froede R., Veltien R. and Salas J.A.
 Nucleic acids which code for the enzyme activities of the spinosyn
 biosynthesis
 Patent: WO 0116303-A 1 08-MAR-2001;

JOURNAL
 BAYER AG (DE)
 Location/Qualifiers
 1..50000
 /organism="Saccharopolyspora spinosa"
 /mol_type="unassigned DNA"
 /db_xref="taxon:60894"

ORIGIN

Query Match 100.0%; Score 828; DB 6; Length 50000;
 Best Local Similarity 100.0%; Pred. No.0; Mismatches 0; Gaps 0;
 Matches 828; Conservative 0;

QY 1 GTGTTCCAGTGGCCGACCAACATGCAAGATTGGGAGATGTATGACTGTGACG 60
 DB 44797 GTGTTCCAGTGGCCGACCAACATGCAAGATTGGGAGATGTATGACTGTGACG 44856
 QY 61 CCGTTCTGAATCGGTTCGGCGCGGCCCTCTGCGCATCCACAGGCTATCTGGAGAAC 120
 DB 44857 CCGTTCTGAATCGGTTCGGCGCGGCCCTCTGCGCATCCACAGGCTATCTGGAGAAC 44916
 QY 121 GAGGGCGGGCTTCCTGGAGAGGCGCGGACCGGCTGACCGACCTGTGCGCGAAGG 180
 DB 44917 GAGGGCGGGCTTCCTGGAGAGGCGCGGACCGGCTGACCGACCTGTGCGCGAAGG 44976
 QY 181 ACCGTCTCGATGGCGCGCTTCGACTGCTGATGTGGGGTCCGGTACCGACAACAGCG 240
 DB 44977 ACCGTCTCGATGGCGCGCTTCGACTGCTGATGTGGGGTCCGGTACCGACAACAGCG 45036
 QY 241 CTGCGCGTTCGGCGCGCAACCGGATCCAGATCAACCGGCTACCGCTCAGCGGTCGCA 300
 DB 45037 CTGCGCGTTCGGCGCGCAACCGGATCCAGATCAACCGGCTACCGCTCAGCGGTCGCA 45096
 QY 301 GTGGCCATCGCCGCTGATGGCGAGCGGAACGGGACTAGACCGGCGGTGACTCTCG 360
 DB 45097 GTGGCCATCGCCGCTGATGGCGAGCGGAACGGGACTAGACCGGCGGTGACTCTCG 45156
 QY 361 TGGTGCATGCATGTCCTGCGCGTACCCGGAATGCTTCGACGCGCGCTGCGCATG 420
 DB 45157 TGGTGCATGCATGTCCTGCGCGTACCCGGAATGCTTCGACGCGCGCTGCGCATG 45216
 QY 421 CAGTGCCTGTGAGATGTCGGAACCGGACCGTGGCCATCCGGGAAATCTTTCAGATGTC 480
 DB 45217 CAGTGCCTGTGAGATGTCGGAACCGGACCGTGGCCATCCGGGAAATCTTTCAGATGTC 45276
 QY 481 AAACCCGGTGGCATCTCCGCGCTGACCGAGTGTCTAAAGAGGCGGCGGGAGT 540
 DB 45277 AAACCCGGTGGCATCTCCGCGCTGACCGAGTGTCTAAAGAGGCGGCGGGAGT 45336
 QY 541 CCGGTGTCGGGGAGCAGGTGGCGACCGGCTTCGATTCGATTCGCTGTCGCAACTTCTG 600
 DB 45337 CCGGTGTCGGGGAGCAGGTGGCGACCGGCTTCGATTCGATTCGCTGTCGCAACTTCTG 45396
 QY 601 GAATCGCTGTGAGCGGGGGTTCGAGATCTTCGATTGGAGAGCGTGTGTCGAGAGCC 660
 DB 45397 GAATCGCTGTGAGCGGGGGTTCGAGATCTTCGATTGGAGAGCGTGTGTCGAGAGCC 45456
 QY 661 CGGACTTCATGCGCGGAGTTTCGCGAAGAGTGTGCTGCGGACCGAGATCGCGAG 720
 DB 45457 CGGACTTCATGCGCGGAGTTTCGCGAAGAGTGTGCTGCGGACCGAGATCGCGAG 45516
 QY 721 AGGATACGGCGCGGCTGTGCGCGCTGGGCGCGCGGCTGTGCGATTATGAGAAATATGCC 780
 DB 45517 AGGATACGGCGCGGCTGTGCGCGCTGGGCGCGCGGCTGTGCGATTATGAGAAATATGCC 45576

QY 781 CACGACATGGGCTATGCGATTCTGACGGCGCGGAGACCGGTCGCTGA 828
 DB 45577 CACGACATGGGCTATGCGATTCTGACGGCGCGGAGACCGGTCGCTGA 45624

RESULT 4
 AY007564
 LOCUS
 DEFINITION Saccharopolyspora spinosa probable lysR-type transcriptional
 regulator, probable keto acyl reductase, spinosad biosynthetic gene
 cluster, complete sequence, and probable exodeoxyribonuclease V
 genes, complete cds, and unknown gene.

ACCESSION
 AY007564
 GI:13162633

VERSION
 AY007564.1

KEYWORDS
 SOURCE Saccharopolyspora spinosa
 ORGANISM Saccharopolyspora spinosa
 Bacteria; Actinobacteridae; Actinomycetales;
 Pseudonocardineae; Pseudonocardaceae; Saccharopolyspora.

REFERENCE 1 (bases 1 to 80161)
 Waldron, C., Matsushima, P., Roestek, P. R. Jr., Broughton, M. C.,
 Turner, J., Madduri, K., Crawford, K. P., Merlo, D. J. and Baltz, R. H.
 Cloning and analysis of the spinosad biosynthetic gene cluster of
 Saccharopolyspora spinosa
 Chem. Biol. 8 (5), 487-499 (2001)

JOURNAL
 MEDLINE
 21257765

PUBMED
 11358695

REFERENCE 2 (bases 1 to 80161)
 Waldron, C., Matsushima, P., Roestek, P. R. Jr., Broughton, M. C.,
 Turner, J., Madduri, K., Crawford, K. P., Merlo, D. J. and Baltz, R. H.
 Direct Submission
 Submitted (18-AUG-2000) Dow Agrosciences, 9330 Zionsville Rd.,
 Indianapolis, IN 46268, USA

TITLE
 JOURNAL
 Indianapolis, IN 46268, USA

FEATURES
 source
 1..80161
 Location/Qualifiers
 /organism="Saccharopolyspora spinosa"
 /mol_type="genomic DNA"
 /db_xref="taxon:60894"
 complement(1135..1971)
 /note="ORF L16"
 /codon_start=1
 /transl_table=11
 /product="probable lysR-type transcriptional regulator"
 /protein_id="AAC23282.1"
 /db_xref="GI:13162633"
 /translation="METRELRYFAVAEELHFGRAAQRIGIAQPPUSRTIADLEQRLG
 VILQRTSRKVSILTEAGAMLITEGRIILGALAAARRTORAATSOPLVLAAGAGSG
 ELLAKLIDAYAEPGAIVADILCESQPKTLHDGADVALIHOPEDPYAEIDIELN
 TEOOVAIILPTSHPLASEPHVMADVSLPDLPLAMPBGDGVYPPGPGVENVQTLF
 QMIALGRITVWNPBESSRVNLEGLAAVPLDAPDVTTVIAMPPHSRSLAGLVRYAT
 L"
 2024..2791
 /note="ORF L15"
 /codon_start=1
 /transl_table=11
 /product="probable keto acyl reductase"
 /protein_id="AAC23281.1"
 /db_xref="GI:13162653"
 /translation="MVLVRRFRFRATLESNEQTIALVTGANKIGIYEIAAGCALGWS
 VGIAGARDHQRGSDAVAKLRADQVDAFAVSLDVTDAVAAALAEERAGRADVLVNA
 AGIAGAMPBEPSTVPASRAVENIVGVETVAMLEPLRSEPRPIVNSQSHVAS
 LTLQTPRGVDLIGIGAVSRKTFPLNATITQYAKELSDPINIKINNACPGYVATDLNGF
 HGTSTPADGARITAIRLATIPDDGPTGCMDDDANNVW"
 complement(3416..4165)
 /gene="spnS"
 /complement(3416..4165)
 /note="probably involved in fosfomycin biosynthesis"
 /gene="spnS"
 /codon_start=1
 /evidence="not experimental"
 /transl_table=11
 /product="probable N-dimethyltransferase"
 /protein_id="AAC23280.1"

/db_xref="GI:13162652"
 /translation="MSRVSDFPASTSVSPDHADIDVAHSARGRDWAAGEVVOV
 VTRLEPAOSLIDVACGTAHLERRPRAVAKAGLESDAMEETIIRRVPEVPIHIG
 IRDPLGEPDVIITCLCTPAAMRTVVDLRRYTRMAHHLAPGVAVIEPWPFPKFI
 DCFVGAHNGEERVISRUSHSVLEGRTSRMVRYTVAPETGIRDFTEFELUSTED
 EYTALEAGIRAEYLPGAPNGRGLFVGIRN"
 complement (4168. .5325)
 /gene="spnr"
 complement (4168. .5325)
 /gene="spnr"
 complement (4168. .5325)
 /note="probably involved in forosamine biosynthesis"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="probable aminotransferase"
 /protein_id="AAG23279.1"
 /db_xref="GI:13162651"
 /translation="MINLHQPILGTBEBDAIEAVPASNWIGLPRTTPEAFPHLH
 VDEQVPLNSTALPLTVQYLDIGPDVVLPISFVAANALASGAPVPCDV
 PRTLPDLDVARTPATKAVLHYGSGPEVTAIDFCSEKGLMIEDSACAVAS
 SVHGACGFRGDLATSPDMKILVYGGGMFYADPELARARLAYHLEOMSGFD
 SAKSNRWMIDRVEDIGORLIGDMTALGVSQVLRKIPETINPRELITQYDRLSDY
 PCVLPPTLPDGHVSHYFYVQVLAPETRDOVQOMLRGITYSTRYPLKHPVITYRA
 DCKLPASADACRRTLLPLHPSLDDAEVTVADFEQKAVEHHISQSPLRK"
 complement (5363. .6751)
 /gene="spnq"
 complement (5363. .6751)
 /gene="spnq"
 complement (5363. .6751)
 /note="probably involved in forosamine biosynthesis"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="probable NDP-hexose-3,4-dehydratase"
 /protein_id="AAG23278.1"
 /db_xref="GI:13162650"
 /translation="MQSRKTRALGKGARVTSQDDTCAITATMWPDAKORILASVRYD
 HREQSPFVAGSTPIRPSGVALDEDDVALVEALRLIAAGNARFESEFAVFG
 LRKALVNSGSSANLIALSLTSPGLGEARLRPGDEVITAAGVPTTINPELVQGLV
 VFVDELGTYNATPDRIKAATERTRAIMLANTGNPPADEIAIEAKHEHLPVEDN
 CDVAGSTRGLTGRFDLTYSFYPAHHITSGEGCVLTGSLRLARIIESLRQMGD
 CKCEPVNTCRKRPDYHGLTLPGYDHKTPSHGVNLKTTDLOALALQSLTSA
 FGSAARRWRRLRBSGLPGILLVATPHSDPSFGFATISADAGFTRALVNFLE
 VSNIGTRLPFGNITRHPAFEQVRYRIADALTNDSIYDRTFWGVYGTIDQMLDYV
 VESIAFVAKSS"
 complement (7083. .8450)
 /gene="spnp"
 complement (7083. .8450)
 /note="probably involved in forosamine addition"
 /codon_start=1
 /transl_table=11
 /product="probable NDP-forsosamy1transferase"
 /protein_id="AAG23277.1"
 /db_xref="GI:13162649"
 /translation="MILGMLPGCSIAIGFMRVLEFTPLPASHFENVLVLAMALPAAG
 HEVRVAICPNMYSMTGAGLTAVPVGDLDLILSLAKNELVYSGSVDEBGRHPELF
 DELSINGRDVDAEQALVDDRLDMEGFAEKMOPLVVMIDAMVCSGACVAVALG
 ARHVMLVALDVSGLRSGFLYEOSKRPBOVDPIGLAKLAKGATGDEBRLVIG
 CATIDPISKMRLLPYLDYISMRFYPYNGPVLVPMBAERPKRVCVCTRLTKRRLS
 RTVEQYGSODEQAMERLIRGAARLDVEVYATISDSEVRMGBELSVNVEHETPA
 NELLESVLIHHGTTQETATVNGPOLILPGFWDESRRAELIADRGAQLVADPA
 FTTEDVDVGOULARLDEPSFANANALIRREIESSPHDIYPRLEKLVABERNRTQG
 SDGHP"
 complement (8967. .10427)
 /gene="spno"
 complement (8967. .10427)
 /note="involved in forosamine biosynthesis"
 /codon_start=1
 /transl_table=11
 /product="probable NDP-hexose-2,3-dehydratase"
 /protein_id="AAG23276.1"
 /db_xref="GI:13162648"

/translation="MSSSVENAEASAPALGSMNTRFVDSALSACNGMIPTEPHCWL
 ADRGENSPENRIRPDRLSKWKCDASTENLVHADGEPFTEGLOVENTNCAASMOQ
 PNIQAOGLICGLVKEIRIDVTLHCLMAKEMPGVNNVLTCLSPYQARTNSYTAHRS
 VPRTVYDFLGRGRVLYVDLQSQSGWFRKRRNMVYEQEVPVLPDPCNLTUQ
 VLAIRKDNITVMKOTRTVLSTCPHDSATGEBLASEPFOAVARSLSHGIDSSIS
 EAVWFEBKARVRLRATRVPLSVDSKRYRDTETIAQDGKYEFAVIVASATREVA
 SMTQPMIEPRECEGIALLVKRIIGVILHGLVHARAEAGYKMAEIAIPVQGSVANYOST
 PSNDMPPELDVLTADDEIVYESIIEEGGRFYQAOQRVRIIEVHDFAPRPSDBR
 MWTIQGELIRSTHPLNIQARSLVASLHSLMGLR"
 complement (10436. .11434)
 /gene="spnq"
 complement (10436. .11434)
 /note="involved in forosamine biosynthesis"
 /codon_start=1
 /transl_table=11
 /product="probable NDP-hexose-3-ketoreductase"
 /protein_id="AAG23275.1"
 /db_xref="GI:13162647"
 /translation="MRKPRVIGVLCASFAMRWMLPAMCDVAETEVAVAVARDPAKAE
 RFARPGENVIGYQRLREPDIDAVVVPPLPGMHAEMIGALAEADKHLAEKPLTTP
 ASDRILVGLARKNLLRENYLFLHGRGHDVVDLQSGEIGREFTVFGIPLP
 DTDIRTELGGLDLDTGYPARANHFLGPIITVYASSHAEQSGVDISGSVLIQ
 SEGGTVAHLGFEVHAYRSAVELWSSRGRLVWDAFTPPAEWQAVIRIERKGVDEL
 LPAEDQVRKAVTAFARDIRAGTGVDDPAVAGDSGSMIOQALVEALGOARCGST"
 complement (11530. .12492)
 /gene="spnm"
 complement (11530. .12492)
 /note="involved in spinosyn aglycone biosynthesis"
 /codon_start=1
 /transl_table=11
 /product="Spnm"
 /protein_id="AAG23274.1"
 /db_xref="GI:13162646"
 /translation="WENAVSGTVLVENIPMPREDRPIITRAVGTGHLGSGVAPSYLR
 TGEPEPTELAVALDRGMVAVITDYBEGLPSTHTYVGAQGMIDAAARQRLPG
 SGLTTPDVGIVGQAQGSAPFAGEHPTVAELRTRAAGAVPIDLDIDITHRNG
 VFTSPDVAIVGHAAVYDLPFDLTEAGRTAVDVBELGAPELVTRFGRLESDY
 DTSLFPOPRMARLALASVAGRPVPLVYHSTDELTVPFAGERLDSYRAACT
 PYRNHPLSGLAHPAALASSRVVYSWDEHFSBSAISGRDAR"
 complement (12696. .13547)
 /gene="spnl"
 complement (12696. .13547)
 /note="involved in spinosyn aglycone biosynthesis; possibly involved in polyketide bridging"
 /codon_start=1
 /transl_table=11
 /product="methyltransferase-like protein"
 /protein_id="AAG23273.1"
 /db_xref="GI:13162645"
 /translation="MESIPDALAGRPRLHHGVYAGYREDAGATPMSDADOLTLDFI
 DKALRFGAHLFDLGCNGOPVVRACASGVRVTGTVNQHILAAATRLANETLAS
 LEFDLVDAQLPYDGFQOAMQSVQVVDQAAALREVRHLLPEGRVVLGDIITR
 VLPDEEVAALWTGTATTLSFTLALVEAEFEILEVLDLPAQTRCMWSWVDELRLL
 DELAGFPAVAVGTQGYLDIAAKHGPBAQLAIVAEYRKPDYARNESMGFMILL
 OARXQGS"
 complement (13592. .14785)
 /gene="spnk"
 complement (13592. .14785)

Query Match 100.0%; Score 828; DB 1; Length 80161;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GTGTTGCCAGGTGGCGACCAACATCGACAGGTTGGGACAGATGTATGACTGTGCAAG 60
 DB 20168 GTGTTGCCAGGTGGCGACCAACATCGACAGGTTGGGACAGATGTATGACTGTGCAAG 20227
 Oy 61 CCGTTGCTGAACCTGTCGCGGGGCGCCCTCGGCCCATCCACCAACGGCTCTGGGAAAC 120
 DB 20228 CCGTTGCTGAACCTGTCGCGGGGCGCCCTCGGCCCATCCACCAACGGCTCTGGGAAAC 20287

QY 121 GACGGGCGGGCTTCTCTGACGAGACGCCGACCCGCTACCGACCTTGTCCCGAAACG 180
 DB 20288 GACGGGCGGGCTTCTCTGACGAGAGCGCGGACCCGCTACCGACCTTGTCCCGAAACG 20347
 QY 181 ACCGCTCTGAGATGGCGCGCTTGAAGTGTGATGGGGGTGGGTTACCGGAAACAACAGCG 240
 DB 20348 ACCGCTCTGAGATGGCGCGCTTGAAGTGTGATGGGGGTGGGTTACCGGAAACAACAGCG 20407
 QY 241 CTGGCGCTGCGCGGACCAACCGGATCCAGATCACCGGATCACCGTCCAGCAGGTGCA 300
 DB 20408 CTGGCGCTGCGCGGACCAACCGGATCCAGATCACCGGATCACCGTCCAGCAGGTGCA 20467
 QY 301 GTGGCCATGCGCGCTGATGGCGGACCGGAACTAAGCCACCGGGTGAATTCTTG 360
 DB 20468 GTGGCCATGCGCGCTGATGGCGGACCGGAACTAAGCCACCGGGTGAATTCTTG 20527
 QY 361 TGGCTGATGAGATGCTCCGCTGACCGGAACTAAGCCACCGGGTGAATTCTTG 420
 DB 20528 TGGCTGATGAGATGCTCCGCTGACCGGAACTAAGCCACCGGGTGAATTCTTG 20587
 QY 421 CAGTCGCTGTTGAGATGCTCCGACCGGACCGGATCCGCGGAAATCCTTGAGTACTC 480
 DB 20588 CAGTCGCTGTTGAGATGCTCCGACCGGACCGGATCCGCGGAAATCCTTGAGTACTC 20647
 QY 481 AAACCCGGTGGCATCTCTCGCGCTCACCGAGTGTCAAACGAGAACCGGGCGGGGATG 540
 DB 20648 AAACCCGGTGGCATCTCTCGCGCTCACCGAGTGTCAAACGAGAACCGGGCGGGGATG 20707
 QY 541 CCGGTGTCCGGGGACAGGTGGCGGACCGGCTTGAGATGCTGCTGAGCAACTTCTG 600
 DB 20708 CCGGTGTCCGGGGACAGGTGGCGGACCGGCTTGAGATGCTGCTGAGCAACTTCTG 20767
 QY 601 GAATGCTGCTGACGAGCGGGGTTGAGATCTTCGATTGGAGGACGCTGTCTGAGAGAC 660
 DB 20768 GAATGCTGCTGACGAGCGGGGTTGAGATCTTCGATTGGAGGACGCTGTCTGAGAGAC 20827
 QY 661 CCGTACTTCATGCGGAGTTCGCGAAGAGCTGCTGCGCACCAAGACGGGATCGCGAC 720
 DB 20828 CCGTACTTCATGCGGAGTTCGCGAAGAGCTGCTGCGCACCAAGACGGGATCGCGAC 20887
 QY 721 AGGTACGGGCGCGGCTGTCCCGGCTGGGCGCGCGCTGCGATTATGAAATATGCC 780
 DB 20888 AGGTACGGGCGCGGCTGTCCCGGCTGGGCGCGCGCTGCGATTATGAAATATGCC 20947
 QY 781 CACGACATGGGCTATGCGATTCTGACGCGCGGAAACCGGTCGCTGA 828
 DB 20948 CACGACATGGGCTATGCGATTCTGACGCGCGGAAACCGGTCGCTGA 20995

RESULT 5
 ARI65018 80161 bp DNA linear PAT 17-OCT-2001
 LOCUS ARI65018
 DEFINITION Sequence 1 from patent US 6274350.
 ACCESSION ARI65018
 VERSION ARI65018.1 GI:16238344
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE 1 (bases 1 to 80161)
 AUTHORS Baltz,R.H., Crawford,K.P., Broughton,M.,Christine., Madduri,K.,
 Merlo,D.J., Turner,J.R., Treadway,P.J., and Waldron,C.
 TITLE Biosynthetic genes for spinosyn insecticide production
 JOURNAL Patent: US 6274350-A 1 14-AUG-2001;
 FEATURES
 Source Location/Qualifiers
 1..80161
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN
 Query Match 100.0%; Score 828; DB 6; Length 80161;
 Best Local Similarity 100.0%; Pred. No. 0;

Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTGTTGCCAGTGGGCGACCAATCCGAGAGGTTGGGCGAGATGATGACCTGTGACG 60
 DB 20168 GTGTTGCCAGTGGGCGACCAATCCGAGAGGTTGGGCGAGATGATGACCTGTGACG 20227
 QY 61 CCGTGTGAACTCGGATCGCGGCGGCGGCTTCCGCTGCGCATCACAGCGCTACTGAGAGAC 120
 DB 20228 CCGTGTGAACTCGGATCGCGGCGGCGGCTTCCGCTGCGCATCACAGCGCTACTGAGAGAC 20287
 QY 121 GACGGGCGGGCTTCTCTGACGAGACCGGCGACCGGCTTCAACCGACTTGTCCCGAAACG 180
 DB 20288 GACGGGCGGGCTTCTCTGACGAGACCGGCGACCGGCTTCAACCGACTTGTCCCGAAACG 20347
 QY 181 ACCGCTCTGAGATGGCGCGCTTGAAGTGTGATGGGGGTGGGTTACCGGAAACAACAGCG 240
 DB 20348 ACCGCTCTGAGATGGCGCGCTTGAAGTGTGATGGGGGTGGGTTACCGGAAACAACAGCG 20407
 QY 241 CTGGCGCTGCGCGGACCAACCGGATCCAGATCACCGGATCACCGTCCAGCAGGTGCA 300
 DB 20408 CTGGCGCTGCGCGGACCAACCGGATCCAGATCACCGGATCACCGTCCAGCAGGTGCA 20467
 QY 301 GTGGCCATGCGCGCTGATGGCGGACCGGAACTAAGCCACCGGGTGAATTCTTG 360
 DB 20468 GTGGCCATGCGCGCTGATGGCGGACCGGAACTAAGCCACCGGGTGAATTCTTG 20527
 QY 361 TGGCTGATGAGATGCTCCGCTGACCGGAACTAAGCCACCGGGTGAATTCTTG 420
 DB 20528 TGGCTGATGAGATGCTCCGCTGACCGGAACTAAGCCACCGGGTGAATTCTTG 20587
 QY 421 CAGTCGCTGTTGAGATGCTCCGACCGGACCGGATCCGCGGAAATCCTTGAGTACTC 480
 DB 20588 CAGTCGCTGTTGAGATGCTCCGACCGGACCGGATCCGCGGAAATCCTTGAGTACTC 20647
 QY 481 AAACCCGGTGGCATCTCTCGCGCTCACCGAGTGTCAAACGAGAACCGGGCGGGGATG 540
 DB 20648 AAACCCGGTGGCATCTCTCGCGCTCACCGAGTGTCAAACGAGAACCGGGCGGGGATG 20707
 QY 541 CCGGTGTCCGGGGACAGGTGGCGGACCGGCTTGAGATGCTGCTGAGCAACTTCTG 600
 DB 20708 CCGGTGTCCGGGGACAGGTGGCGGACCGGCTTGAGATGCTGCTGAGCAACTTCTG 20767
 QY 601 GAATGCTGCTGACGAGCGGGGTTGAGATCTTCGATTGGAGGACGCTGTCTGAGAGAC 660
 DB 20768 GAATGCTGCTGACGAGCGGGGTTGAGATCTTCGATTGGAGGACGCTGTCTGAGAGAC 20827
 QY 661 CCGTACTTCATGCGGAGTTCGCGAAGAGCTGCTGCGCACCAAGACGGGATCGCGAC 720
 DB 20828 CCGTACTTCATGCGGAGTTCGCGAAGAGCTGCTGCGCACCAAGACGGGATCGCGAC 20887
 QY 721 AGGTACGGGCGCGGCTGTCCCGGCTGGGCGCGCGCTGCGATTATGAAATATGCC 780
 DB 20888 AGGTACGGGCGCGGCTGTCCCGGCTGGGCGCGCGCTGCGATTATGAAATATGCC 20947
 QY 781 CACGACATGGGCTATGCGATTCTGACGCGCGGAAACCGGTCGCTGA 828
 DB 20948 CACGACATGGGCTATGCGATTCTGACGCGCGGAAACCGGTCGCTGA 20995

RESULT 6
 BD137649 80161 bp DNA linear PAT 18-SEP-2002
 LOCUS BD137649
 DEFINITION Biosynthetic genes for spinosyn insecticide production.
 ACCESSION BD137649
 VERSION BD137649.1 GI:23232594
 KEYWORDS JP 2002505881-A/1.
 SOURCE Saccharopolyspora spinosa
 ORGANISM Saccharopolyspora spinosa
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Pseudonocardiales; Pseudonocardiales; Saccharopolyspora.
 REFERENCE 1 (bases 1 to 80161)
 AUTHORS Baltz,R.H., Broughton,M.C., Crawford,K.P., Madduri,K., Merlo,D.J.,
 Treadway,P.J., Turner,J.R. and Waldron,C.

TITLE Biosynthetic genes for spinosyn insecticide production
JOURNAL Patent: JP 2002505881-A 1 26-FEB-2002;
COMMENT OS Saccharopolyspora spinosa
PN JP 2002505881-A/1
PD 26-FEB-2002
PF 16-FEB-1998 JP 2000535754
PR 09-MAR-1998 US 09/036987
PI RICHARD H BALTZ, M CHRISTINE BROUGHTON, KATHRYN P CRAWFORD, PI KRISHNAMURTHY MADHURI, DONALD J MERLO, PATTI J TREADWAY, JUAN R PI TURNER
PI CLIVE WALDRON
PC C12N15/09, C12N11/15, C12N11/19, C12N11/21, C12N5/10, C12P19/62, C12Q1/PC 68//
PC C07K14/195, C12N15/00, C12N5/00
CC Biosynthetic genes for spinosyn insecticide production FH
Key Location/Qualifiers
FT source 1..80161
FEATURES Location/Qualifiers
source 1..80161
/organism="Saccharopolyspora spinosa"
/mol_type="genomic DNA"
/db_xref="taxon:60894"
ORIGIN
Query Match 100.0%; Score 828; DB 6; Length 80161;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGTTGCCAGGTGGCGCACCAATCGACAGAGTTGGGACAGATGTATGACTGTGTACG 60
Db 20168 GTGTTGCCAGGTGGCGCACCAATCGACAGAGTTGGGACAGATGTATGACTGTGTACG 20227
QY 61 CCGTTGCTGAACCTGGTGGCGGGGGCCCTGGCGCATCCACCAAGGCTACTGGGAGAAC 120
Db 20228 CCGTTGCTGAACCTGGTGGCGGGGGCCCTGGCGCATCCACCAAGGCTACTGGGAGAAC 20287
QY 121 GACGGGCGGGGCTTCTTGGCAGCAGGCGCGCCGACCGGCTCACCGACCTTGTGCCCGAACGG 180
Db 20288 GACGGGCGGGGCTTCTTGGCAGCAGGCGCGCCGACCGGCTCACCGACCTTGTGCCCGAACGG 20347
QY 181 ACCGTGCTCATGTGGCGGGGTTGACTGCTCGATGTGGGGTGGGTAACCGGACACACGCG 240
Db 20348 ACCGTGCTCATGTGGCGGGGTTGACTGCTCGATGTGGGGTGGGTAACCGGACACACGCG 20407
QY 241 CTGGCGCGTGGCGCGGACCAAGCGCATCCAGATCAACGGGATCAACGTCAGCAGGCTGGAA 300
Db 20408 CTGGCGCGTGGCGCGGACCAAGCGCATCCAGATCAACGGGATCAACGTCAGCAGGCTGGAA 20467
QY 301 GTGGCCATCGCGCTGATTTGGCGCAGCGCAACCGGACTTAAGCCACCGGGTGGACTTCTCG 360
Db 20468 GTGGCCATCGCGCTGATTTGGCGCAGCGCAACCGGACTTAAGCCACCGGGTGGACTTCTCG 20527
QY 361 TGGCTCATGTCCATGTCCCTGCGCGTAACCGGACATGCTTTTGAGCGCGGCTTGGGCCCATG 420
Db 20528 TGGCTCATGTCCATGTCCCTGCGCGTAACCGGACATGCTTTTGAGCGCGGCTTGGGCCCATG 20587
QY 421 CAGTGCCTGTTGAGAGATGTCGGAACCGGACCGGTCATCCGGGAAATCTTTCGAGTATCTC 480
Db 20588 CAGTGCCTGTTGAGAGATGTCGGAACCGGACCGGTCATCCGGGAAATCTTTCGAGTATCTC 20647
QY 481 AAACCCGGTGGCATCTCGCGCTCAACGAGGTGTCMAACGAGAACGCGGCGCGCGAGTGG 540
Db 20648 AAACCCGGTGGCATCTCGCGCGCTCAACGAGGTGTCMAACGAGAACGCGGCGCGCGAGTGG 20707
QY 541 CCGGTGTCGGGAGACAGGTGGCGGACCGGCTTCCGATCTGCTGGCTGAGCACTTCTG 600
Db 20708 CCGGTGTCGGGAGACAGGTGGCGGACCGGCTTCCGATCTGCTGGCTGAGCACTTCTG 20767
QY 601 GAATGCGTGGCTGACGCGGGGGTTGAGATTCCTCGATTGGGAGAGACGTGTCTGCGAGGACC 660
Db 20768 GAATGCGTGGCTGACGCGGGGGTTGAGATTCCTCGATTGGGAGAGACGTGTCTGCGAGGACC 20827

QY 661 CGGTACTTCATGCCGACAGTTGGCGGAGAGCTCCGTGGCACACGACGCGGATCGCGAC 720
Db 20828 CGGTACTTCATGCCGACAGTTGGCGGAGAGCTCCGTGGCACACGACGCGGATCGCGAC 20887
QY 721 AGGTACGGGCGGCTGTCTGCGCGGCTGGCGCGCCCGGCTCTCCGATTTAGAAATATGCC 780
Db 20888 AGGTACGGGCGGCTGTCTGCGCGGCTGGCGCGCCCGGCTCTCCGATTTAGAAATATGCC 20947
QY 781 CACGACATGGGCTATGAGATTCTGACGGCGGCGGAAGCCGGTCGAGTGA 828
Db 20948 CACGACATGGGCTATGAGATTCTGACGGCGGCGGAAGCCGGTCGAGTGA 20995
RESULT 7
LOCUS AR281866 80161 bp DNA linear PAT 10-APR-2003
DEFINITION AR281866 Sequence 1 from patent US 6521406.
ACCESSION AR281866
VERSION AR281866.1 GI:29717767
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unknown.
AUTHORS 1 (bases 1 to 80161)
Baltz, R.H., Broughton, M.C., Crawford, K.P., Madduri, K., Merlo, D.J., Treadway, P.J., Turner, J.R. and Waldron, C.
TITLE Spng, a gene for spinosyn insecticide biosynthesis
JOURNAL Patent: US 6521406-A 1 18-FEB-2003;
FEATURES Location/Qualifiers
source 1..80161
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 828; DB 6; Length 80161;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGTTGCCAGGTGGCGCACCAATCGACAGAGTTGGGACAGATGTATGACTGTGTACG 60
Db 20168 GTGTTGCCAGGTGGCGCACCAATCGACAGAGTTGGGACAGATGTATGACTGTGTACG 20227
QY 61 CCGTTGCTGAACCTGGTGGCGGGGCTTGGCGCATCCACCAAGGCTACTGGGAGAAC 120
Db 20228 CCGTTGCTGAACCTGGTGGCGGGGCTTGGCGCATCCACCAAGGCTACTGGGAGAAC 20287
QY 121 GACGGGCGGGGCTTCTTGGCAGCAGGCGCGCCGACCGGCTCACCGACCTTGTGCCCGAACGG 180
Db 20288 GACGGGCGGGGCTTCTTGGCAGCAGGCGCGCCGACCGGCTCACCGACCTTGTGCCCGAACGG 20347
QY 181 ACCGTGCTCATGTGGCGGGGTTGACTGCTCGATGTGGGGTGGGTAACCGGACACACGCG 240
Db 20348 ACCGTGCTCATGTGGCGGGGTTGACTGCTCGATGTGGGGTGGGTAACCGGACACACGCG 20407
QY 241 CTGGCGCGTGGCGCGGACCAAGCGCATCCAGATCAACGGGATCAACGTCAGCAGGCTGGAA 300
Db 20408 CTGGCGCGTGGCGCGGACCAAGCGCATCCAGATCAACGGGATCAACGTCAGCAGGCTGGAA 20467
QY 301 GTGGCCATCGCGCTGATTTGGCGCAGCGCAACCGGACTTAAGCCACCGGGTGGACTTCTCG 360
Db 20468 GTGGCCATCGCGCTGATTTGGCGCAGCGCAACCGGACTTAAGCCACCGGGTGGACTTCTCG 20527
QY 361 TGGCTCATGTCCATGTCCCTGCGCGTAACCGGACATGCTTTTGAGCGCGGCTTGGGCCCATG 420
Db 20528 TGGCTCATGTCCATGTCCCTGCGCGTAACCGGACATGCTTTTGAGCGCGGCTTGGGCCCATG 20587
QY 421 CAGTGCCTGTTGAGAGATGTCGGAACCGGACCGGTCATCCGGGAAATCTTTCGAGTATCTC 480
Db 20588 CAGTGCCTGTTGAGAGATGTCGGAACCGGACCGGTCATCCGGGAAATCTTTCGAGTATCTC 20647
QY 481 AAACCCGGTGGCATCTCGCGCTCAACGAGGTGTCMAACGAGAACGCGGCGCGCGAGTGG 540

Db 20648 AACCCTGGGATCTCTGCGCTCAACCGAGTCTGTCAACGAGAAAGCGGCGCGGATG 20707
Qy 541 CCGGTGTCCGGGAGACAGTGGCCGACCGGCTTCGGATTCGCTGGCTTACGAACTTCTG 600
Db 20708 CCGGTGTCCGGGAGACAGTGGCCGACCGGCTTCGGATTCGCTGGCTTACGAACTTCTG 20767
Qy 601 GAATGCTGCGGACGCGGCGGTTTCGAGATCTCGATTGGGAGACGCTGCTCGAGAGCC 660
Db 20768 GAATGCTGCGGACGCGGCGGTTTCGAGATCTCGATTGGGAGACGCTGCTCGAGAGCC 20827
Qy 661 CCGTACTTCATCCGACAGTTCGCGGAGAGCTCGCTGCGCACACAGCAGCGAGATCGCGAC 720
Db 20828 CCGTACTTCATCCGACAGTTCGCGGAGAGCTCGCTGCGCACACAGCAGCGAGATCGCGAC 20887
Qy 721 AGGTACGGGCGCGGCTGTCCCGCTGGCGCGCTGCGCTGCGATTATGAAATATGCG 780
Db 20888 AGGTACGGGCGCGGCTGTCCCGCTGGCGCGCTGCGCTGCGATTATGAAATATGCG 20947
Qy 781 CACGACATGGGCTATGCGATTGACAGCGGCGGAGCGGATCGGCTGA 828
Db 20948 CACGACATGGGCTATGCGATTGACAGCGGCGGAGCGGATCGGCTGA 20995

RESULT 8
AX600587/c 36538 bp DNA linear PAT 14-FEB-2003
LOCUS AX600587
DEFINITION Sequence 2 from Patent WO02079477.
ACCESSION AX600587
VERSION AX600587.1 GI:28400303
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Hahn, D.R., Jackson, J.D., Bullard, B.S., Gustafson, G.D., Waldron, C. and Mitchell, J.C.
TITLE Biosynthetic genes for butenyl-epinosyn insecticide production
JOURNAL Patent: WO 02079477-A 2 10-OCT-2002;
Dow Agrosciences LLC (US)
FEATURES
source 1. 36538
/organism="Saccharopolyspora sp. NRRL 30141"
/mol_type="unassigned DNA"
/db_xref="taxon:221471"
/note="LW107129"

ORIGIN
Query Match 6.9%; Score 57; DB 6; Length 36538;
Best Local Similarity 100.0%; Pred. No. 4,6e-18;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 73 TCGGTGCGGCGGCGGCGGCTGCGCATCCACACGAGCTACTGAGAGAGCGGCGG 129
Db 866 TCGGTGCGGCGGCGGCGGCTGCGCATCCACACGAGCTACTGAGAGAGCGGCGG 810

RESULT 9
AX600621/c 22 bp DNA linear PAT 14-FEB-2003
LOCUS AX600621
DEFINITION Sequence 36 from Patent WO02079477.
ACCESSION AX600621
VERSION AX600621.1 GI:28400307
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Hahn, D.R., Jackson, J.D., Bullard, B.S., Gustafson, G.D., Waldron, C. and Mitchell, J.C.
TITLE Biosynthetic genes for butenyl-epinosyn insecticide production
JOURNAL Patent: WO 02079477-A 36 10-OCT-2002;

FEATURES
source 1. 22
Location/Qualifiers
/organism="Saccharopolyspora spinosa"
/mol_type="unassigned DNA"
/db_xref="taxon:60894"

ORIGIN
Query Match 2.7%; Score 22; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 769 GAGAAATATGCCACGACATGG 790
Db 22 GAGAAATATGCCACGACATGG 1

RESULT 10
E07847/c 457 bp DNA linear PAT 29-SEP-1997
LOCUS E07847
DEFINITION DNA sequence participating the biosynthesis of avermectin.
ACCESSION E07847
VERSION E07847.1 GI:2175980
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Omura, S. and Ikeda, H.
TITLE DIRECT FERMENTATION OF 5-OXO-AVERMECTIN DERIVATIVE AND MICROORGANISM FOR ITS PRODUCTION
JOURNAL Patent: JP 1994189774-A 2 12-JUL-1994;
KITASATO INST:THE
COMMENT
OS Streptomyces avermitilis
PN JP 1994189774-A/2
PD 12-JUL-1994
PF 24-DEC-1992 JP 1992344605
PI OMURA SATOSHI, IKEDA HARUO
PC C12N15/53, C12N1/21, C12P19/62, (C12N1/21, C12R1:465), (C12P19/62, C12R1:465);
PC C12R1:465);
CC strandedness: Double;
CC topology: Linear;
FH key Location/Qualifiers
FH source 1. 457
FT /organism="Streptomyces avermitilis" FT
/strain="K2062".
FEATURES
source 1. 457
Location/Qualifiers
/organism="Streptomyces avermitilis"
/mol_type="genomic DNA"
/db_xref="taxon:33903"

ORIGIN
Query Match 2.7%; Score 22; DB 6; Length 457;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 143 AGGCGCGGACCGGCTCACCGA 164
Db 94 AGGCGCGGACCGGCTCACCGA 73

RESULT 11
E07848/c 457 bp DNA linear PAT 29-SEP-1997
LOCUS E07848
DEFINITION DNA sequence participating the biosynthesis of avermectin.
ACCESSION E07848
VERSION E07848.1 GI:2175981
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Hahn, D.R., Jackson, J.D., Bullard, B.S., Gustafson, G.D., Waldron, C. and Mitchell, J.C.
TITLE Biosynthetic genes for butenyl-epinosyn insecticide production
JOURNAL Patent: WO 02079477-A 36 10-OCT-2002;

REFERENCE Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
1 (bases 1 to 457)
AUTHORS Omura,S. and Ikeda,H.
TITLE DIRECT FERMENTATION OF 5-OXO-AVERMECTIN DERIVATIVE AND MICROORGANISM FOR ITS PRODUCTION
JOURNAL Patent: JP 1994189774-A 3 12-JUL-1994;
KITASATO INST.THE
COMMENT OS Streptomycetes avermilitilis
PN JP 1994189774-A/3
PD 12-JUL-1994
PF 24-DEC-1992 JP 1992344605
PI OMURA SATOSHI, IKEDA HARUO
PC C12N15/53, C12N1/21, C12P19/62, (C12N1/21, C12R1:465), (C12P19/62, C12R1:465);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1. .457
/strain='K2082'.
FEATURES location/Qualifiers
source 1. .457
/organism='Streptomyces avermilitilis'
/mol_type='genomic DNA'
/db_xref='taxon:33903'
ORIGIN
Query Match 2.7%; Score 22; DB 6; Length 457;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 143 AGCGCGCGACCGGCTCACCGA 164
|||||
94 AGCGCGCGACCGGCTCACCGA 73
DB
RESULT 12
E07846/c 1410 bp DNA linear PAT 29-SEP-1997
LOCUS E07846
DEFINITION DNA sequence participating the biosynthesis of avermectin.
ACCESSION E07846
VERSION E07846.1 GI:2175979
KEYWORDS JP 1994189774-A/1.
SOURCE Streptomycetes avermilitilis
ORGANISM Streptomycetes avermilitilis
Bacteria; Actinobacterii; Actinobacteridae; Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
1 (bases 1 to 1410)
Omura,S. and Ikeda,H.
TITLE DIRECT FERMENTATION OF 5-OXO-AVERMECTIN DERIVATIVE AND MICROORGANISM FOR ITS PRODUCTION
JOURNAL Patent: JP 1994189774-A 1 12-JUL-1994;
KITASATO INST.THE
COMMENT OS Streptomycetes avermilitilis
PN JP 1994189774-A/1
PD 12-JUL-1994
PF 24-DEC-1992 JP 1992344605
PI OMURA SATOSHI, IKEDA HARUO
PC C12N15/53, C12N1/21, C12P19/62, (C12N1/21, C12R1:465), (C12P19/62, C12R1:465);
CC strandedness: Double;
CC topology: Linear;
FH Key location/Qualifiers
FT source 1. .1410
/organism='Streptomyces avermilitilis' FT
/strain='ATCC31271'.
FEATURES location/Qualifiers
source 1. .1410
/organism='Streptomyces avermilitilis'
/mol_type='genomic DNA'

ORIGIN /db_xref='taxon:33903'
Query Match 2.7%; Score 22; DB 6; Length 1410;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 143 AGCGCGCGACCGGCTCACCGA 164
|||||
720 AGCGCGCGACCGGCTCACCGA 699
DB
RESULT 13
AB032524/c 5998 bp DNA linear BCT 23-SEP-1999
LOCUS AB032524
DEFINITION Streptomycetes avermilitilis avermectin biosynthetic gene cluster (aver, aveF, aveD), complete cds.
ACCESSION AB032524
VERSION AB032524.1 GI:5921164
KEYWORDS C5-O-methyltransferase, C-5 ketoreductase.
SOURCE Streptomycetes avermilitilis
ORGANISM Streptomycetes avermilitilis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
1 (bases 1 to 5998)
Ikeda,H., Nonomura,T., Usami,M., Ohta,T. and Omura,S.
TITLE Organization of the biosynthetic gene cluster for the polyketide antihelminthic macroide avermectin in Streptomyces avermilitilis
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9509-9514 (1999)
MEDLINE 99380548
PUBMED 10449723
REFERENCE 2 (bases 1 to 5998)
AUTHORS Ikeda,H., Nonomura,T., Usami,M., Ohta,T. and Omura,S.
TITLE Direct Submision
JOURNAL Submitted (20-SEP-1999) Haruo Ikeda, School of Pharmaceutical Sciences, Kitasato University, Microbial Chemistry, 5-9-1 Shirokane, Minato-ku, Tokyo 108-8641, Japan
(E-mail: ikeda@mc.pharm.kitasato-u.ac.jp, Tel: +81-3-5791-6242, Fax: +81-3-3444-6197)
FEATURES location/Qualifiers
source 1. .5998
/organism='Streptomyces avermilitilis'
/mol_type='genomic DNA'
/db_xref='taxon:33903'
gene 327. .3176
/gene='aver'
327. .3176
/gene='aver'
/note='contains helix-turn-helix DNA binding domain positive response regulator.'
/codon_start=1
/transl_table=1
/protein_id='BA84600.1'
/db_xref='GI:5921165'
/translation='WQGVSCIHPPKREBEITLVDFETOPALRLALTECAAGYKLV
AEGMGCGKSTFLGEALHTAASGFALVRAAGPADRQPLGVQQLNDPAPEDTAR
TAVPMPVOHVRGALERLAAGAPLAIIDVOADGSLCKMLTHTSPTSLILC
TALACSPADPVLEAEIMROTAPEFRIITLDCISLGVGLSDRCARPTAPADYCLT
VTGGNPLLRLALLEHSEADAPRPAEASALPPQAPPPVGVGRGVSYIATL
SRITTAIRQTRAGALAVUGRABADLPQLGLASASVSTRGLRLEAGLTSGFRBP
VAEPAALDADPDSRRALHRAALQHHDAAPDVNARLPAARHAGPAAVSVLRA
AEQSLADQDVASVSCLEAVGACVRRERPEIRIRLAAAFGRNTNVAEHLADVA
TLREGELTGHOTLALVPLVNHGRLGARBARMLNADARGLCADGSPMAWPBS
TAHLAARADPAARDPGRDPADKPELPPOSGAPORPADRGQCPPTAALMALPAG
TSBAARAAROVRRSPILTSTVLVNAVRIILARTGVYTAIDWCHRLGEATRRRC
PGNOAHILAVRAELSCRGILADAKCEQALTHVPHSHSVAPAGGLAOULACTAM
GRIDEATQULSHPEPEALPHSVGLGLVRLRAGHFLMWNLPAVRLPTLRGRVAREM
GLDHPVLLPWRTAAEAFRLIGETKRADQQLTDLQVSPHSGNPVVRGTALRLQTA
PAERLRLISLPSRGTDRGPDGRALAAALADLGAAYHSRNEPVRASATVRAMQAKC
GAQALCDLSLPSRGTDRGPDGRALAAALADLGAAYHSRNEPVRASATVRAMQAKC
VSTVEQHLTVYRKLTIRRELPTRLRLHADQN'
complement (3223. .4131)
/gene='aveF'


```

CDS
    complement(3223..4131)
    /gene="aveF"
    /note="post-polyketide modification"
    /codon_start=1
    /transl_table=11
    /product="C-5 ketoreductase"
    /protein_id="BAA84601.1"
    /db_xref="GI:5921166"
    /translation="MRECGRGIPVAVTRKVESLITSGPCLLSEKGNFTVADLKRAA
    LVTGASRGIGRAIAQRLAEGLVAVHGSQKAAAEVTEALERVGSEAFVRAQLR
    DVAVDLTFEVENRLDGRPLVAVNNAVGSPAPAGAGKGDVTSVGVAAEFD
    LERINVRAPFIIQRLALRIGEGRVINVSALTRIAMPLIPYAMS KGLAEIARLA
    LEIGARGITVTVAPGITATDNNRAMRDVPGVERISALTALRRLGRPDIGVVAFL
    ASDEGRWVTGQLVDSAGGLCTAPAMS"
    complement(4911..5762)
    /gene="aveD"
    /complement(4911..5762)
    /gene="aved"
    /note="post-polyketide modification"
    /codon_start=1
    /transl_table=11
    /product="CS-O-methyltransferase"
    /protein_id="BAA84602.1"
    /db_xref="GI:5921167"
    /translation="MMAVPEQTPPSLEVDYVDRLTDIMNRALGNTHTAGVWPQBD
    GSPGKADRLTDLIGKRGITTRVLDVGGSGKPAYRLASAPVDVGVTVSVQ
    VELATLAKQSHVADRVFTRADAMELPPDPSFDAMALBCLHWPSPQVTRAR
    VLRPGRLAVTQVALRAFGRTMKRGECSTQLAVPALVHIDVYAGMDAGLEIHEL
    TDIGQVGPSPFALRDHYNEHLDEYAAAFGIVAMERKVAQCTTLPTWIDIGYVVL
    TARRPGE"

ORIGIN
Query Match      2.7%; Score 22; DB 1; Length 5998;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      143 AGCGCCGCGACCGGCTCACC GA 164
      |||||||
DB      5614 AGCGCCGCGACCGGCTCACC GA 5593

RESULT 14
BD217469/c      12381 bp      DNA      linear      PAT 17-JUL-2003
LOCUS      Polyketides, their preparation, and materials for use therein.
ACCESSION      BD217469
VERSION      BD217469.1 GI:33027239
KEYWORDS      JP 2002519066-A/1
SOURCE      Streptomyces avermitilis
ORGANISM      Streptomyces avermitilis
      Bacteria; Actinobacteridae; Actinomycetales;
      Streptomycineae; Streptomycetaceae; Streptomyces.
      1 (bases 1 to 12381)
AUTHORS      Kellenberger, J.L., Leadlay, P.F., Staunton, J., Engwall, K.J.S. and
      McArthur, H.A.1.
TITLE      Polyketides, their preparation, and materials for use therein
JOURNAL      Patent: JP 2002519066-A 1 02-JUL-2002;
      BIOTICA TECHNOLOGY LTD, PFIZER INC
COMMENT      OS Streptomyces avermitilis
      PN JP 2002519066-A/1
      PD 02-JUL-2002 JP 2000558217
      PR 06-JUL-1998 GB 9814622.8
      PI JOHANNES LAURENZ KELLENBERGER, PETER FRANCIS LEADLAY, JAMES PI
      STAUNTON,
      PT KIM JONELLE STUTZMAN ENGMALL, HAMISH ALASTAIR IRVINE MCARTHUR
      PC C12N15/09, C07K19/00, C12N1/21, C12N9/02, C12N9/04, C12P17/06, PC
      C12P19/62//
      CC (C12N1/21, C12N1:465), C12N15/00
      FH Polyketides, their preparation, and materials for use therein
      FT Key Location/Qualifiers
      1..12381
      FT source
      /organism="Streptomyces avermitilis".
  
```

```

FEATURES
    source
    Location/Qualifiers
    1..12381
    /organism="Streptomyces avermitilis"
    /mol_type="genomic DNA"
    /db_xref="taxon:33903"

ORIGIN
Query Match      2.7%; Score 22; DB 6; Length 12381;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      143 AGCGCCGCGACCGGCTCACC GA 164
      |||||||
DB      333 AGCGCCGCGACCGGCTCACC GA 312

RESULT 15
AX006889/c      12381 bp      DNA      linear      PAT 06-SEP-2000
LOCUS      Sequence 1 from Patent WO0001827.
ACCESSION      AX006889
VERSION      AX006889.1 GI:9994904
KEYWORDS
SOURCE      Streptomyces avermitilis
ORGANISM      Streptomyces avermitilis
      Bacteria; Actinobacteridae; Actinomycetales;
      Streptomycineae; Streptomycetaceae; Streptomyces.
      1
AUTHORS      Kellenberger, J.L., Leadlay, P.F., Staunton, J., McArthur, H.A. and
      Stutzman-Engwall, K.J.
TITLE      Polyketides, their preparation, and materials for use therein
JOURNAL      Patent: WO 0001827-A 1 13-JAN-2000;
      KELLENBERGER JOHANNES LAURENZ (CH); LEADLAY PETER FRANCIS (GB);
      STAUNTON JAMES (GB); BIOTICA TECH LTD (GB); PFIZER (US); MCARTHUR
      HAMISH ALASTAIR IRVINE (US); STUTZMAN ENGMALL KIM JONELLE (US)
FEATURES
    source
    Location/Qualifiers
    1..12381
    /organism="Streptomyces avermitilis"
    /mol_type="unassigned DNA"
    /db_xref="taxon:33903"

Search completed: August 29, 2005, 18:38:21
Job time : 3897 secs
  
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: August 29, 2005, 15:06:10 ; Search time 585 Seconds
(without alignments)
8378.711 Million cell updates/sec

Title: US-10-069-353A-7

Sequence: 1 gtgtctgcacaggtcgcacac...cgcggaagcaggtctcggtcga 828

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

```
Minimum DB seq length: 0
```

Maximum DB seq length: 20000000000

Post-processing: Listing first 45 summaries

Database : N_GeneSeq_16Dec04:*

```

1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query				ID	Description
No.	Score	Match	Length	DB		
1	828	100.0	828	4	AAf68318	AAf68318 S. spinos
C	2	828	100.0	45624	4	AAf68315 S. spinos
3	828	100.0	50000	4	AAf68312	AAf68312 S. spinos
4	828	100.0	80161	2	AAz21501	AAz21501 DNA fragm
C	5	57	6.9	36538	10	ABv75558 Saccharo
C	6	22	2.7	22	10	ABv75562 Saccharo
C	7	22	2.7	457	2	AAQ70316 Oxidoredu
C	8	22	2.7	457	2	AAQ70315 Oxidoredu
C	9	22	2.7	1410	2	AAQ80524 Oxidoredu
C	10	22	2.7	14381	3	AAz58381 Streptomy
C	11	20	2.4	20	10	ABv75561 Saccharo
12	20	2.4	237	12	ADp62503	ADp62503 Maize ca
13	20	2.4	263	12	ADp62374	ADp62374 Maize ca
14	20	2.4	290	12	ADp62375	ADp62375 Maize ca
15	20	2.4	294	12	ADp62373	ADp62373 Maize ca
16	20	2.4	305	12	ADp62372	ADp62372 Maize ca
17	20	2.4	433	12	ADp62502	ADp62502 Maize ca
18	20	2.4	1654	13	ADs58180	ADs58180 Bacteri
19	20	2.4	1650	13	ADs4549	ADs4549 Zsa maye
20	2.4	1832	6	AAI45918	AAI45918 Terridact	

C	21	20	2	4	2805	8	ACA40862
C	22	20	2	4	2985	8	ACA38551
C	23	20	2	4	3107	2	AAV68082
C	24	20	2	4	8051	2	AAT72685
C	25	20	2	4	8160	2	AAK25774
C	26	20	2	4	110000	4	AAI99682_40
C	27	20	2	4	110000	4	AAI99683_40
C	28	19	2	3	909	5	AAH94478
C	29	19	2	3	1050	5	AAH94372
C	30	19	2	3	1188	10	ADC36141
C	31	19	2	3	1418	10	ADC36144
C	32	19	2	3	1418	10	ADC36202
C	33	19	2	3	4741	4	ABL27984
C	34	19	2	3	110000	2	AAK19980_03
C	35	19	2	3	110000	2	AAK19980_04
C	36	19	2	3	273254	3	AAAC1914
C	37	18	2	2	265	10	ABX84638
C	38	18	2	2	268	7	ADST70438
C	39	18	2	2	272	10	ADST70078
C	40	18	2	2	273	7	ADST70078
C	41	18	2	2	275	7	ADST70280
C	42	18	2	2	276	7	ADST70495
C	43	18	2	2	276	7	ADST70414
C	44	18	2	2	285	7	ADST70463
C	45	18	2	2	295	6	ABL73622
ALIGNMENTS							
RESULT 1							
ID	AAF88318	AAF88318 standard; DNA; 828 BP.					
XX	AAF88318;						
DT	28-AUG-2001	(first entry)					
DE	S. spinosa	DNA fragment encoding ORF1, SEQ ID 7.					
XX							
KW	Forosamine; trimethylxanthose; polyketide synthase; biosynthesis;						
KM	spinovyn; polyketide aglycone; transgenic plant; insect resistance;						
KV	macrolide; insecticidal; O-methyltransferase; ds.						
XX							
OS	Saccharopolyspora spinosa.						
XX							
PN	DE19957268-A1.						
PD	08-MAR-2001.						
PF	29-NOV-1999;	99DE-01057268.					
XX							
PR	27-AUG-1999;	99DE-01040596.					
XX							
PA	(PARB) BAYER AG.						
PI	Eberz G,	Moehrl V,	Froede R,	Velten R,	Salas JA;		
XX							
DR	WPI; 2001-267102/28.						
XX	P-PSDB; AAB70948.						
PT	New nucleic acid encoding enzymes for spinovyn biosynthesis, useful for						
XX	recombinant production of insecticidal spinovyns and their derivatives.						
PS	Claim 7a; Page 102-104; 354pp; German.						
CC	This invention describes a novel method nucleic acid (i) and its encoded						
CC	polypeptide (ii) containing at least one region that encodes an enzymatic						
CC	activity involved in biosynthesis of spinovyn. (i) are used (i) to						
CC	identify, inactivate or modulate genes involved in the biosynthesis of						
CC	(ii); (iii) to generate a library of polyketide synthases; (iii) for						
CC	adding forosamine or trimethylxanthose to a spinovyn; polyketide						
CC	aglycone; and (iv) for recombinant production of the corresponding						

enzymes, which are used for production of (II), their precursors or derivatives, including production of transgenic plants that express (II) and thus have increased resistance to insects. (I) are also useful as markers for sequencing of the Saccharopolyspora spinosa genome. (II) are macrolides with insecticidal, but not antibacterial, activity, and can also be used to raise specific antibodies, useful for identifying expression clones in a gene bank. Cells transformed with (I) may produce CC (II) at significantly increased levels or produce new derivatives of CC (II). This sequence, ORF 1, encodes an S. spinosa O-methyltransferase protein

XX Sequence 828 BP; 140 A; 262 C; 285 G; 141 T; 0 U; 0 Other;

Query Match 100.0%; Score 828; DB 4; Length 828;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GTGTTGCCAGTGGCGCACCAACATCGACAGTTGGGAGATGATGACTGTGACG 60
DB 1 GTGTTGCCAGTGGCGCACCAACATCGACAGTTGGGAGATGATGACTGTGACG 60
QY 61 CCGTTGCTGAACCTCGGTGCGGGGCGCCCTGCGCATCAGCGGCTACTGGGAGAAC 120
DB 61 CCGTTGCTGAACCTCGGTGCGGGGCGCCCTGCGCATCAGCGGCTACTGGGAGAAC 120
QY 121 GACGGGCGGGCTTCTCTGGCAGAGCGCCGCGACCGGCTTCACGACCTTGTCCGCGAACGG 180
DB 121 GACGGGCGGGCTTCTCTGGCAGAGCGCCGCGACCGGCTTCACGACCTTGTCCGCGAACGG 180
QY 181 ACCGTCCTGATGAGCGGCGCTTGTGCTGATGATGGGGTGGGTTACCGGACACCGACG 240
DB 181 ACCGTCCTGATGAGCGGCGCTTGTGCTGATGATGGGGTGGGTTACCGGACACCGACG 240
QY 241 CCGCGGCTGGCGGAGACGCGATCCAGATCACCGGATCAGCGTCAACGAGGTGACAA 300
DB 241 CCGCGGCTGGCGGAGACGCGATCCAGATCACCGGATCAGCGTCAACGAGGTGACAA 300
QY 301 GTGGCCATGCGCGCTGATGTCGCGACGCGACCGGACCTAAGCCACCGGCTGACTTCTCG 360
DB 301 GTGGCCATGCGCGCTGATGTCGCGACGCGACCGGACCTAAGCCACCGGCTGACTTCTCG 360
QY 361 TCGCGTGAATGCAATGCTTCCGCGTACCGGACCAATGCTTTGACGCGCGCTGGGCCATG 420
DB 361 TCGCGTGAATGCAATGCTTCCGCGTACCGGACCAATGCTTTGACGCGCGCTGGGCCATG 420
QY 421 CAGTGGCTTGGAGATGTCGAGACCGGACCGGTCATCCGGGAAATCCTTGAGTATCTC 480
DB 421 CAGTGGCTTGGAGATGTCGAGACCGGACCGGTCATCCGGGAAATCCTTGAGTATCTC 480
QY 481 AAACCCGCGTGCATCTCTCGCGCTCAACGAGTGTCAAAAGAAACGCGGCGCGGATG 540
DB 481 AAACCCGCGTGCATCTCTCGCGCTCAACGAGTGTCAAAAGAAACGCGGCGCGGATG 540
QY 541 CCGGTGTCTCGGGGAGACAGTGGGCGGACCGGCTTCGGATGCTGCTGGTGAAGCACTTCTG 600
DB 541 CCGGTGTCTCGGGGAGACAGTGGGCGGACCGGCTTCGGATGCTGCTGGTGAAGCACTTCTG 600
QY 601 GAATGCTGCTGTCGACGCGGGGTTTCGAGATCTCGATTGGGAGAGAGTGTGTCGAGAAC 660
DB 601 GAATGCTGCTGTCGACGCGGGGTTTCGAGATCTCGATTGGGAGAGAGTGTGTCGAGAAC 660
QY 661 CCGTACTTTCATGCGGAGTTTCGCGAAGAGTCTGCTGCGCACACAGCGAGATTCGCGAC 720
DB 661 CCGTACTTTCATGCGGAGTTTCGCGAAGAGTCTGCTGCGCACACAGCGAGATTCGCGAC 720
QY 721 AGGTACGGGGCGGCTGTGCGCGGCTGGGCGCGCGGCTTGGGATTAAGAAATATATGCC 780
DB 721 AGGTACGGGGCGGCTGTGCGCGGCTGGGCGCGCGGCTTGGGATTAAGAAATATATGCC 780
QY 781 CACGACATGGGCTATGCGATTTCTGACGCGCGGAAAGCGGCTGAGCTGA 828
DB 781 CACGACATGGGCTATGCGATTTCTGACGCGCGGAAAGCGGCTGAGCTGA 828

```

RESULT 2
AAF88315/C

ID AAF88315 standard; DNA; 45624 BP.

XX AAF88315;

AC 28-AUG-2001 (first entry)

XX S. spinosa DNA fragment SEQ ID 4.

XX Forosamine; trimethylrhamsone; polyketide synthase; biosynthesis;

KW spinosyn; polyketide aglycone; transgenic plant; insect resistance;

XX macrolide; insecticidal; ds.

OS Saccharopolyspora spinosa.

XX DE1957268-A1.

XX 08-MAR-2001.

XX 29-NOV-1999; 99DE-01057268.

XX 27-AUG-1999; 99DE-01040596.

XX (FARB) BAYER AG.

XX Eberz G, Moehrl V, Froede R, Velten R, Salas JA;

XX WPI, 2001-267102/28.

XX New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for recombinant production of insecticidal spinosyns and their derivatives.

XX Claim 7; Page 58-74; 354pp; German.

XX This invention describes a novel method nucleic acid (I) and its encoded polypeptide (II) containing at least one region that encodes an enzymatic activity involved in biosynthesis of spinosyns. (I) are used (i) to identify, inactivate or modulate genes involved in the biosynthesis of (II); (ii) to generate a library of polyketide synthases; (iii) for adding forosamine or trimethylrhamsone to a spinosyn or polyketide aglycone; and (iv) for recombinant production of the corresponding enzymes, which are used for production of (II), their precursors or derivatives, including production of transgenic plants that express (II) and thus have increased resistance to insects. (I) are also useful as markers for sequencing of the Saccharopolyspora spinosa genome. (II) are macrolides with insecticidal, but not antibacterial, activity, and can also be used to raise specific antibodies, useful for identifying expression clones in a gene bank. Cells transformed with (I) may produce CC (II) at significantly increased levels or produce new derivatives of CC (II). This sequence represents a genomic DNA fragment of the S. spinosa genome which contains the coding regions for proteins involved in forosamine and trimethylrhamsone biosynthesis

XX Sequence 45624 BP; 7933 A; 14913 C; 14940 G; 7838 T; 0 U; 0 Other;

Query Match 100.0%; Score 828; DB 4; Length 45624;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GTGTTGCCAGTGGCGCACCAACATCGACAGTTGGGAGATGATGACTGTGACG 60
DB 828 GTGTTGCCAGTGGCGCACCAACATCGACAGTTGGGAGATGATGACTGTGACG 769
QY 61 CCGTTGCTGAACCTCGGTGCGGGGCGCCCTGCGCATCCACCGGCTACTGGGAGAAC 120
DB 768 CCGTTGCTGAACCTCGGTGCGGGGCGCCCTGCGCATCCACCGGCTACTGGGAGAAC 709
QY 121 GACGGGCGGGCTTCTCTGGCAGAGCGCCGCGACCGGCTTCACGACCTTGTGCGCGAACGG 180
DB 708 GACGGGCGGGCTTCTCTGGCAGAGCGCCGCGACCGGCTTCACGACCTTGTGCGCGAACGG 649

```

[illegible]

Query Match	Beat Local Similarity	100.0%;	Score 828;	DB 4;	Length 50000;
Matches	828;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
XX	WPI; 2001-267102/28.				
XX	New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for				
XX	recombinant production of insecticidal spinosyns and their derivatives.				
XX	Claim 7; Page 14-31; 354pp; German.				
CC	This invention describes a novel method nucleic acid (I) and its encoded				
CC	polypeptide (II) containing at least one region that encodes an enzymatic				
CC	activity involved in biosynthesis of spinosyns. (I) are used (i) to				
CC	identify, inactivate or modulate genes involved in the biosynthesis of				
CC	(II); (ii) to generate a library of polypeptide synthetases; (iii) for				
CC	adding forosamine or trimethylrhinamose to a spinosyn or polypeptide				
CC	aglycone; and (iv) for recombinant production of the corresponding				
CC	enzymes, which are used for production of (II), their precursors or				
CC	derivatives, including production of transgenic plants that express (II)				
CC	and thus have increased resistance to insects. (I) are also useful as				
CC	markers for sequencing of the Saccharopolyspora spinosa genome. (II) are				
CC	macrolides with insecticidal, but not antibacterial, activity, and can				
CC	also be used to raise specific antibodies, useful for identifying				
CC	expression clones in a gene bank. Cells transformed with (I) may produce				
CC	(II) at significantly increased levels or produce new derivatives of				
CC	(II). This sequence represents a genomic DNA fragment of the S. spinosa				
CC	genome which contains the coding regions for proteins involved in				
CC	forosamine, trimethylrhinamose and polyketide synthase biosynthesis				
XX					
SQ	Sequence 50000 BP; 8484 A; 16290 C; 16535 G; 8691 T; 0 U; 0 Other;				
Query Match	100.0%;	Score 828;	DB 4;	Length 50000;	
Beat Local Similarity	100.0%;	Pred. No. 0;			
Matches	828;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
QY	1	GTGTTGCCAGGTGGCGCACCAACATCTGGCAGAGCTGGGCAGATGTATGACTCTGGTCAGG	60		
DB	44797	GTGTTGGCAGGTGGCGCACCAACATCTGGCAGAGCTGGGCAGATGTATGACTCTGGTCAGG	44856		
QY	61	CCGTTGTGAACTCCGTCGCGGGGGGGCCCTGTGGCCATCCACACGGCTACTGGGAGAAC	120		
DB	44857	CCGTTGTGAACTCCGTCGCGGGGGGGCCCTGTGGCCATCCACACGGCTACTGGGAGAAC	44916		
QY	121	GACGGGCGGGCTTCTCTGGCAGACGCGCGCGACCTGACCTTGTGCGCGAACG	180		
DB	44917	GACGGGCGGGCTTCTCTGGCAGACGCGCGCGACCTGACCTTGTGCGCGAACG	44976		
QY	181	ACCGTGTCCAGTGGCGGGCGTTTGCATCTGCTGATGTGGGTGGCGTACCGGACCAACGCG	240		
DB	44977	ACCGTGTCCAGTGGCGGGCGTTTGCATCTGCTGATGTGGGTGGCGTACCGGACCAACGCG	45036		
QY	241	CTGGCGCGTCCGCGCGCGACACCGGATCCAGATACCGGCACTCCGTCACGCGAGGCGAAC	300		
DB	45037	CTGGCGCGTCCGCGCGCGACACCGGATCCAGATACCGGCACTCCGTCACGCGAGGCGAAC	45096		
QY	301	GTTGGCAGTCCGCGCTGATTTGGCGACGCGGACCGGAGCTAAAGCACCGGGGTGACTTCTCG	360		
DB	45097	GTTGGCAGTCCGCGCTGATTTGGCGACGCGGACCGGAGCTAAAGCACCGGGGTGACTTCTCG	45156		
QY	361	TGCGTGCATGTCATGTCCCTGCCGTACCCCGACCAATGCTTTGACGCGCGCTGGAGCCATG	420		
DB	45157	TGCGTGCATGTCATGTCCCTGCCGTACCCCGACCAATGCTTTGACGCGCGCTGGAGCCATG	45216		
QY	421	CAGTGCCTGTTTGGAGATGTCCGAACCGGACCTGTGCGCATCCGGGAAATCTCTTGAGATATC	480		
DB	45217	CAGTGCCTGTTTGGAGATGTCCGAACCGGACCTGTGCGCATCCGGGAAATCTCTTGAGATATC	45276		
QY	481	AAACCCGGGTGCATCTCCGCGGTCCACCGAGTGTCTAAACGAGAACCGGGCGCGGAGT	540		
DB	45277	AAACCCGGGTGCATCTCCGCGGTCCACCGAGTGTCTAAACGAGAACCGGGCGCGGAGT	45336		
QY	541	CCGCTGTCCGGGACAGGTGGCGGACCGGCTTTCGATCTGCTGGCTGAGCACTTCTTG	600		
DB	45337	CCGCTGTCCGGGACAGGTGGCGGACCGGCTTTCGATCTGCTGGCTGAGCACTTCTTG	45396		
QY	601	GATTCGCTGCGTGCAGCGGGCTTGAAGATCTCGATTTGGAGGACGCTGTCTGTGAGACG	660		

Db 45397 GAAATGCTGCGTGAAGCGGGGTTGAGATCTCGATTGGAGAGAGTGTCTCGAGGACC 45456
Oy 661 CGGTAATTCATGCGCCGACATTCGCGGAAAGAGCTCGTGCGACACAGCAGGATCGCGAC 720
Db 45457 CGGTAATTCATGCGCCGACATTCGCGGAAAGAGCTCGTGCGACACAGCAGGATCGCGAC 45516
Oy 721 AGTAACGGGCGGCGTGTGCGCGCTGCGGCGCGCGGTCTCGATTATGAGAAATATGCC 780
Db 45517 AGTAACGGGCGGCGTGTGCGCGCTGCGGCGCGCGGTCTCGATTATGAGAAATATGCC 45576
Oy 781 CACGACATGGGCTATGCGATTCTGAACGGCGGAAAGCGGTGCGCTGA 828
Db 45577 CACGACATGGGCTATGCGATTCTGAACGGCGGAAAGCGGTGCGCTGA 45624

RESULT 4
AA221501
ID AA221501 standard; DNA; 80161 BP.
XX
XX AA221501;
AC
XX 01-DEC-1999 (first entry)
XX
XX DNA fragment of Saccharopolyspora spinosa containing biosynthetic genes.
DE
XX
XX Spinosyn biosynthetic enzyme; open reading frame; ORF.
KM insecticidal microides; arachnid; nematode; insect; polyketide;
KM polyketide synthase; PKS; extender module; initiator module;
KM acyl transferase domain; AT; acyl carrier protein; ACP;
KM beta-ketosynthase domain; KS; KR; dehydratase domain; DH;
KM enoyl reductase domain; ER; beta-ketoreductase; insecticide; ss.
XX
OS Saccharopolyspora spinosa.

XX
XX Location/Qualifiers
FH
FT complement(1135. .1971)
FT
FT /*tag= a
FT /product= "ORFL16"
FT /note= "Protein involved in transcription control"
FT 2024. .2791
FT
FT /*tag= b
FT /product= "ORFL15"
FT /note= "Keto acyl reductase"
FT complement(3416. .4165)
FT
FT /*tag= c
FT /product= "spnS"
FT /note= "Spinosyn biosynthesis protein S"
FT complement(4168. .5325)
FT
FT /*tag= d
FT /product= "spnR"
FT /note= "Spinosyn biosynthesis protein R"
FT complement(5363. .6751)
FT
FT /*tag= e
FT /product= "spnQ"
FT /note= "Spinosyn biosynthesis protein Q"
FT 7083. .8450
FT
FT /*tag= f
FT /product= "spnP"
FT /note= "Spinosyn biosynthesis protein P"
FT 8967. .10427
FT
FT /*tag= g
FT /product= "spnO"
FT /note= "Spinosyn biosynthesis protein O"
FT 10436. .11434
FT
FT /*tag= h
FT /product= "spnN"
FT /note= "Spinosyn biosynthesis protein N"
FT complement(11530. .12492)
FT
FT /*tag= i
FT /product= "spnM"
FT /note= "Spinosyn biosynthesis protein M"
FT complement(12696. .13547)
FT
FT CDS

FT
FT /*tag= j
FT /product= "spnL"
FT /note= "Spinosyn biosynthesis protein L"
FT complement(13592. .14785)
FT
FT /*tag= k
FT /product= "spnK"
FT /note= "Spinosyn biosynthesis protein K"
FT complement(14799. .16418)
FT
FT /*tag= l
FT /product= "spnJ"
FT /note= "Spinosyn biosynthesis protein J"
FT 16556. .17743
FT
FT /*tag= m
FT /product= "spnI"
FT /note= "Spinosyn biosynthesis protein I"
FT complement(17749. .18501)
FT
FT /*tag= n
FT /product= "spnH"
FT /note= "Spinosyn biosynthesis protein H"
FT complement(18541. .19713)
FT
FT /*tag= o
FT /product= "spnG"
FT /note= "Spinosyn biosynthesis protein G"
FT 20168. .20995
FT
FT /*tag= p
FT /product= "spnF"
FT /note= "Spinosyn biosynthesis protein F"
FT 21111. .28898
FT
FT /*tag= q
FT /product= "spnA"
FT /note= "Spinosyn biosynthesis protein A"
FT /transl_except= (Pos:26940. .26942, aa:Pro)
FT 21126. .22379
FT
FT /*tag= r
FT /note= "Beta-ketosynthase domain (KS1): part of the
FT initiator module"
FT 22692. .23669
FT
FT /*tag= s
FT /note= "Acyl transferase domain (AT1): part of the
FT initiator module"
FT 23793. .24041
FT
FT /*tag= t
FT /note= "Acyl carrier protein domain (ACP1): part of the
FT initiator module"
FT 24102. .25349
FT
FT /*tag= u
FT /note= "Beta-ketosynthase domain (KS1): part of extender
FT module 1"
FT 25683. .26684
FT
FT /*tag= v
FT /note= "Acyl transferase domain (AT1): part of extender
FT module 1"
FT 27582. .28121
FT
FT /*tag= w
FT /note= "Beta-ketoreductase domain (KR1): part of extender
FT module 1"
FT 28404. .28649
FT
FT /*tag= x
FT /note= "Acyl carrier protein domain (ACP1): part of
FT extender module 1"
FT 29024. .30295
FT
FT /*tag= z
FT /note= "Beta-ketosynthase domain (KS2): part of extender
FT module 2"
FT 30629. .31621
FT
FT /*tag= aa
FT /note= "Acyl transferase domain (AT2): part of extender
FT module 2"
FT 31697. .32254
FT
FT /*tag= ab
FT /note= "Dehydratase domain (DH2): part of extender module
FT 2"
FT 33035. .34072
FT
FT misc_feature

```
PT      /*tag= ac
PT      /note= "Enoyl reductase domain (ER2) : part of extender
PT      module 2"
PT      34082. .34621
PT      /*tag= ad
PT      /note= "Beta-ketoreductase domain (KR2) : part of extender
PT      module 2"
PT      34886. .30295
PT      /*tag= ae
PT      /note= "Acyl carrier protein domain (ACP2) : part of
PT      extender module 2"
PT      35419. .44931
PT      /*tag= af
PT      /product= "spnc"
PT      /note= "Spinobyn biosynthesis protein C"
PT      35518. .36786
PT      /*tag= ag
PT      /note= "Beta-ketosynthase domain (KS3) : part of extender
PT      module 3"
PT      37108. .38097
PT      /*tag= ah
PT      /note= "Acyl transferase domain (AT3) : part of extender
PT      module 3"
PT      38916. .35374
PT      /*tag= y
PT      /product= "spnc"
PT      /note= "Spinobyn biosynthesis protein C"
PT      38992. .39528
PT      /*tag= ai
PT      /note= "Beta-ketoreductase domain (KR3) : part of extender
PT      module 3"
PT      39790. .40035
PT      /*tag= aj
PT      /note= "Acyl carrier protein domain (ACP3) : part of
PT      extender module 3"
PT      40102. .41373
PT      /*tag= ak
PT      /note= "Beta-ketosynthase domain (KS4) : part of extender
PT      module 4"
PT      41713. .42705
PT      /*tag= al
PT      /note= "Acyl transferase domain (AT4) : part of extender
PT      module 4"
PT      43615. .44157
PT      /*tag= am
PT      /note= "Beta-ketoreductase domain (KR4) : part of extender
PT      module 4"
PT      44431. .44676
PT      /*tag= an
PT      /note= "Acyl carrier protein domain (ACP4) : part of
PT      extender module 4"
PT      44966. .59752
PT      /*tag= ao
PT      /product= "spnd"
PT      /note= "Spinobyn biosynthesis protein D"
PT      45077. .46348
PT      /*tag= ap
PT      /note= "Beta-ketosynthase domain (KS5) : part of extender
PT      module 5"
PT      46691. .47674
PT      /*tag= aq
PT      /note= "Acyl transferase domain (AT5) : part of extender
PT      module 5"
PT      47753. .48310
PT      /*tag= ar
PT      /note= "Dehydratase domain (DH5) : part of extender module
PT      5"
PT      49226. .49771
PT      /*tag= as
PT      /note= "Beta-ketoreductase domain (KR5) : part of extender
PT      module 5"
PT      50009. .50254
PT      misc_feature
```

```
Query Match      100.0%; Score 828; DB 2; Length 80161;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1      GTGTTGCCAGGTGGCCGACCAACATGCGAGAGGTTGGCAGATGTATGACCTGTACG 60
Db      20168 GTGTTGCCAGGTGGCCGACCAACATGCGAGAGGTTGGCAGATGTATGACCTGTACG 20227

Qy      61      CCGTGTCTGAATCTGGTCCGGGGGCGCCCTCTGCGCATCAACAGGGCTACTTGGGGAAC 120
Db      20228 CCGTGTCTGAATCTGGTCCGGGGGCGCCCTCTGCGCATCAACAGGGCTACTTGGGGAAC 20287

Qy      121     GACGGGGGGGCTTCCGTGGAGAGAGCGCGGACCCGGCTCAGCGACCTTGTGCGGAACG 180
Db      20288 GACGGGGGGGCTTCCGTGGAGAGAGCGCGGACCCGGCTCAGCGACCTTGTGCGGAACG 20347

Qy      181     ACCGTGCTCGATGCGCGCGCTTGCATCTGATGTGGGGTGGCGTACCGGACAAACGAGG 240
Db      20348 ACCGTGCTCGATGCGCGCGCTTGCATCTGATGTGGGGTGGCGTACCGGACAAACGAGG 20407

Qy      241     CTGCGCGTGGCGCGCGGACAAACCGGATCCAGATCAACCGCATCACCGTCAAGCCAGGTGCA 300
Db      20408 CTGCGCGTGGCGCGCGGACAAACCGGATCCAGATCAACCGCATCACCGTCAAGCCAGGTGCA 20467

Qy      301     GTGGCCATGGCGCGGTGATGGCGAGCGGAAACCGGAGCTAAGCCACCGGGTGAATTCTCG 360
Db      20468 GTGGCCATGGCGCGGTGATGGCGAGCGGAAACCGGAGCTAAGCCACCGGGTGAATTCTCG 20527

Qy      361     TGGGTGATGTCATGTCCTGCGGTACCGCGACAAATCTTTGACCGCGCTGGGCGATG 420
Db      20528 TGGGTGATGTCATGTCCTGCGGTACCGCGACAAATCTTTGACCGCGCTGGGCGATG 20587

Qy      421     CAGTGCCTGTTGGAGATGTCGCCAACCGGACCGTGGCATTCGGGAAATCTTTCAGATATC 480
Db      20588 CAGTGCCTGTTGGAGATGTCGCCAACCGGACCGTGGCATTCGGGAAATCTTTCAGATATC 20647

Qy      481     AAACCCGGTGGCATCTCTCGGCGTCAACGAGTGTGTAACGAGAGCGGGCGGAGATG 540
Db      20648 AAACCCGGTGGCATCTCTCGGCGTCAACGAGTGTGTAACGAGAGCGGGCGGAGATG 20707

Qy      541     CCGGTGTCGGGGGACAGATGGCCGACCGGCGCTTGCAGATCTGCTGCGTGAACAATTCTG 600
Db      20708 CCGGTGTCGGGGGACAGATGGCCGACCGGCGCTTGCAGATCTGCTGCGTGAACAATTCTG 20767

Qy      601     GAATCGCTGCTGTCAGCGCGGGGTTTCAGATCTCTGATTGGAGAGACGTGTCTGAGAGACC 660
Db      20768 GAATCGCTGCTGTCAGCGCGGGGTTTCAGATCTCTGATTGGAGAGACGTGTCTGAGAGACC 20827

Qy      661     CGGTACTTCAATGGCCGAGTTCCGCGAAGCTGGCTGGCGACACGACCGGGATCGGCGAC 720
Db      20828 CGGTACTTCAATGGCCGAGTTCCGCGAAGCTGGCTGGCGACACGAGGATCGGCGAC 20887

Qy      721     AGGTACGGGCGGCGCTGTCGCGGCTGGGCGCGCGCGTCTGCGATTATGAGAAATATGCC 780
Db      20888 AGGTACGGGCGGCGCTGTCGCGGCTGGGCGCGCGCGTCTGCGATTATGAGAAATATGCC 20947

Qy      781     CACGACATGGGCTATGCGATTCTGACGCGCGGGAAGCCGGTGGCTGCA 828
Db      20948 CACGACATGGGCTATGCGATTCTGACGCGCGGGAAGCCGGTGGCTGCA 20995

RESULT 5
ABV75558/c
ID      ABV75558 standard; DNA; 36538 BP.
XX
XX      ABV75558;
AC
DT      22-JAN-2003 (first entry)
XX
XX      Saccharopolyspora butenyl-spinosyn biosynthetic gene cluster 2.
DB
XX      Butenyl; biosynthetic enzyme; PKS; polyketide synthetase; macrolide;
KW      metabolite; spinosyn; gene; ds.
```

```
XX OS Saccharopolyspora sp.
XX Key location/Qualifiers
FH CDS complement(114. .938)
FT /tag= a
FT /product= "bushF"
FT /note= "No start codon given"
FT CDS 1389. .2561
FT /tag= b
FT /product= "bushG"
FT CDS 2601. .3353
FT /tag= c
FT /product= "bushH"
FT CDS complement(3359. .4546)
FT /tag= d
FT /product= "bushI"
FT CDS 4684. .6303
FT /tag= e
FT /product= "bushJ"
FT CDS 6317. .7510
FT /tag= f
FT /product= "bushK"
FT CDS 7553. .8406
FT /tag= g
FT /product= "bushL"
FT CDS 8640. .9572
FT /tag= h
FT /product= "bushM"
FT CDS complement(9668. .10666)
FT /tag= i
FT /product= "bushN"
FT CDS complement(10675. .12135)
FT /tag= j
FT /product= "bushO"
FT CDS complement(12864. .14177)
FT /tag= k
FT /product= "bushP"
FT CDS 14627. .15970
FT /tag= l
FT /product= "bushQ"
FT CDS complement(15978. .20488)
FT /tag= p
FT /product= "ORF LII"
FT CDS complement(20536. .21033)
FT /tag= q
FT /product= "ORF LIII"
FT CDS 21179. .21925
FT /tag= r
FT /product= "ORF LIV"
FT CDS complement(22671. .23453)
FT /tag= s
FT /product= "ORF LVI"
FT CDS complement(23687. .24886)
FT /tag= t
FT /product= "ORF LVII"
FT CDS complement(26177. .26923)
FT /tag= u
```

```
FT /product= "ORF LVIII"
FT /note= "No start codon given"
FT CDS 27646. .28476
FT /tag= v
FT /product= "ORF LIX"
FT /note= "No start codon given"
FT WO200279477-A2.
XX 10-OCT-2002.
XX 28-MAR-2002; 2002WO-US009968.
XX 30-MAR-2001; 2001US-0280175P.
XX (IDMC ) DOW AGROSCIENCES LLC.
XX Hahn DR, Jackson JD, Bullard BS, Gustafson GD, Waldron C;
XX Mitchell JC;
XX WPI; 2003-058434/05.
XX P-P&SDB; ABP57683, ABP57684, ABP57685, ABP57686, ABP57687, ABP57688.
XX DR ABP57689, ABP57690, ABP57691, ABP57692, ABP57693, ABP57694, ABP57695,
XX DR ABP57696, ABP57697, ABP57698, ABP57699, ABP57700, ABP57701, ABP57702,
XX DR ABP57703, ABP57704.
XX New butenyl-spinosyn biosynthetic genes, useful for increasing the
XX production of butenyl-spinosyn insecticidal macrocides, or for changing
XX the metabolites or products produced by spinosyn-producing
XX microorganisms.
XX Claim 2; Page 99-119; 218pp; English.
XX
XX The invention relates to a novel DNA molecule comprising a DNA sequence
XX that encodes a butenyl-spinosyn biosynthetic enzyme, a butenyl-spinosyn
XX PKS (polyketide synthetase) domain, or a spinosyn PKS module. The butenyl-
XX -spinosyn biosynthetic genes are useful for increasing the production of
XX butenyl-spinosyn insecticidal macrocides. The genes are also useful for
XX changing the metabolites or products produced by spinosyn-producing
XX microorganisms. The present sequence represents a DNA molecule encoding
XX butenyl-spinosyn biosynthetic enzymes
XX
XX Sequence 36538 BP; 6867 A; 12266 C; 11182 G; 6223 T; 0 U; 0 Other;
SQ
Query Match 6.9%; Score 57; DB 10; Length 36538;
Best Local Similarity 100.0%; Pred. No. 1.9e-17;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 73 TCGGTCGGCGGGCGCCCTGCGCCATCCACACGCGCTACTGGAGAAAGACGGCGG 129
DB 866 TCGGTCGGCGGGCGCCCTGCGCCATCCACACGCGCTACTGGAGAAAGACGGCGG 810
RESULT 6
ABV75562/c
ID ABV75562 standard; DNA; 22 BP.
XX
XX ABV75562;
AC
XX
XX 22-JAN-2003 (first entry)
XX
XX Saccharopolyspora spnF reverse PCR primer.
DE
XX
XX Butenyl; biosynthetic enzyme; PKS; polyketide synthetase; macrocides;
KW metabolite; spinosyn; PCR; primer; ss; spnF.
XX
XX Saccharopolyspora spinosa.
OS
XX
XX WO200279477-A2.
XX
XX 10-OCT-2002.
XX
XX 28-MAR-2002; 2002WO-US009968.
XX
XX 28-MAR-2002; 2002WO-US009968.
```



```

XX 30-MAR-2001; 2001US-0280175P.
PR (DOWC ) DOW AGROSCIENCES LLC.
XX
XX Hahn DR, Jackson JD, Bullard BS, Gustafson GD, Waldron C;
PI Mitchell JC;
XX WPI; 2003-058434/05.
XX
XX New butenyl-spinosyn biosynthetic genes, useful for increasing the
PT production of butenyl-spinosyn insecticidal macrolides, or for changing
PT the metabolites or products produced by spinosyn-producing
PT microorganisms.
XX
XX Example 1; Page 218; 218pp; English.
XX
XX The invention relates to a novel DNA molecule comprising a DNA sequence
CC that encodes a butenyl-spinosyn biosynthetic enzyme, a butenyl-spinosyn
CC PKS (polyketide synthetase) domain, or a spinosyn PKS module. The butenyl
CC -spinosyn biosynthetic genes are useful for increasing the production of
CC butenyl-spinosyn insecticidal macrolides. The genes are also useful for
CC changing the metabolites or products produced by spinosyn-producing
CC microorganisms. The present sequence represents a PCR primer used to
CC amplify the spinosyn biosynthetic gene spnF from a S. spinosa
CC biosynthetic gene cluster
XX
XX Sequence 22 BP; 3 A; 6 C; 5 G; 8 T; 0 U; 0 Other;
SQ
XX
XX Query Match 2.7%; Score 22; DB 10; Length 22;
XX Best Local Similarity 100.0%; Pred. No. 3.5;
XX Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 769 GAGAAATATGCCACGACATCG 790
DB 22 GAGAAATATGCCACGACATCG 1

RESULT 7
AAQ70316/c
ID AAQ70316 standard; DNA; 457 BP.
XX
XX AAQ70316;
XX
XX 10-APR-1995 (first entry)
XX
XX Oxidoreducing avermectin DNA from S. avermectilis FERM P-13155.
DE
XX Modified; oxidoreducing; avermectin; S. avermectilis; wild type; disabled;
XX de.
XX Streptomyces avermectilis.
XX OS
XX JP06189774-A.
XX PN
XX 12-JUL-1994.
XX PD
XX 24-DEC-1992; 92JP-00344605.
XX PF
XX 24-DEC-1992; 92JP-00344605.
XX PR
XX (KITA ) KITASATO KENKYUSHO SH.
XX PA
XX (KITA ) KITASATO KENKYUSHO SH.
XX PP
XX WPI; 1994-259554/32.
XX
XX A new modified DNA fragment encoding an avermectin deriv. - useful for
PT the produ. of the deriv.
XX
XX Claim 5; Page 2; 17pp; Japanese.
XX
XX The sequences in AAQ80524 and AAQ70315-16 represent modified
CC oxidoreducing avermectin genes from S. avermectilis strains deposited as
CC ATCC 31271, FERM P-13154 and FERM P-13155. These genes were introduced
CC ATCC 31271, FERM P-13154 and FERM P-13155. These genes were introduced
CC

```

```

CC into S. avermectilis strains in which the wild type avermectin gene has
CC been disabled such that the modified avermectins are produced in large
CC amounts
XX
XX Sequence 457 BP; 83 A; 145 C; 145 G; 83 T; 0 U; 0 Other;
SQ
XX
XX Query Match 2.7%; Score 22; DB 2; Length 457;
XX Best Local Similarity 100.0%; Pred. No. 2.8;
XX Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 143 AGGCCGCCGACCGGCTCACCGA 164
DB 94 AGGCCGCCGACCGGCTCACCGA 73

RESULT 8
AAQ70315/c
ID AAQ70315 standard; DNA; 457 BP.
XX
XX AAQ70315;
XX
XX 10-APR-1995 (first entry)
XX
XX Oxidoreducing avermectin DNA from S. avermectilis FERM P-13154.
DE
XX Modified; oxidoreducing; avermectin; S. avermectilis; wild type; disabled;
XX de.
XX Streptomyces avermectilis.
XX OS
XX JP06189774-A.
XX PN
XX 12-JUL-1994.
XX PD
XX 24-DEC-1992; 92JP-00344605.
XX PF
XX 24-DEC-1992; 92JP-00344605.
XX PR
XX (KITA ) KITASATO KENKYUSHO SH.
XX PA
XX (KITA ) KITASATO KENKYUSHO SH.
XX PP
XX WPI; 1994-259554/32.
XX
XX A new modified DNA fragment encoding an avermectin deriv. - useful for
PT the produ. of the deriv.
XX
XX Claim 4; Page 2; 17pp; Japanese.
XX
XX The sequences in AAQ80524 and AAQ70315-16 represent modified
CC oxidoreducing avermectin genes from S. avermectilis strains deposited as
CC ATCC 31271, FERM P-13154 and FERM P-13155. These genes were introduced
CC into S. avermectilis strains in which the wild type avermectin gene has
CC been disabled such that the modified avermectins are produced in large
CC amounts
XX
XX Sequence 457 BP; 83 A; 145 C; 145 G; 84 T; 0 U; 0 Other;
SQ
XX
XX Query Match 2.7%; Score 22; DB 2; Length 457;
XX Best Local Similarity 100.0%; Pred. No. 2.8;
XX Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 143 AGGCCGCCGACCGGCTCACCGA 164
DB 94 AGGCCGCCGACCGGCTCACCGA 73

RESULT 9
AAQ80524/c
ID AAQ80524 standard; DNA; 1410 BP.
XX
XX AAQ80524;
XX
XX 10-APR-1995 (first entry)
XX

```

```
DE Oxidoreducing avermectin DNA from S. avermectilis ATCC31271.
XX Modified; oxidoreducing; avermectin; S. avermectilis; wild type; disabled;
KM ds.
XX Streptomyces avermectilis.
OS
XX Key Location/Qualifiers
FH misc_feature 627..1075
FT /tag= a
FT /note= "Avermectin coding sequence"
XX
XX JP06189774-A.
XX 12-JUL-1994.
XX
XX 24-DEC-1992; 92UP-00344605.
XX
XX 24-DEC-1992; 92UP-00344605.
XX
XX (KIRYA ) KIRYASATO KENKYUSHO SH.
XX
XX WPI; 1994-259554/32.
XX
XX A new modified DNA fragment encoding an avermectin deriv. - useful for
PT the prodn. of the deriv.
XX
XX Claim 2; Page 2; 17pp; Japanese.
XX
XX The sequences in AA080524 and AA070315-16 represent modified
CC oxidoreducing avermectin genes from S. avermectilis strains deposited as
CC ATCC 31271, FERM P-13154 and FERM P-13155. These genes were introduced
CC into S. avermectilis strains in which the wild type avermectin gene has
CC been disabled such that the modified avermectins are produced in large
CC amounts
XX
XX Sequence 1410 BP; 232 A; 488 C; 463 G; 227 T; 0 U; 0 Other;
SQ
Query Match 2.7%; Score 22; DB 2; Length 1410;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 143 AGCGCGCGACCGGCTCACC GA 164
DB 720 AGCGCGCGACCGGCTCACC GA 639
RESULT 10
AA258381/c
ID AA258381 standard; DNA; 12381 BP.
XX
XX AA258381;
AC
XX
XX 23-MAY-2000 (first entry)
DT
XX Streptomyces avermectilis avermectin polyketide synthase modules 1+2.
DE
XX Polyketide synthase; avermectin; insecticide; ss.
XX
XX Streptomyces avermectilis.
OS
XX WO200001827-A2.
XX
XX 13-JAN-2000.
XX
XX 06-JUL-1999; 99WO-GB002158.
XX
XX 06-JUL-1998; 98GB-00014622.
XX
XX (BIOT-) BIOTICA TECHNOLOGY LTD.
XX (PF12 ) PFIZER INC.
XX
XX Kellenberger JL, Leadlay PF, Staunton J, Stutzman-Engwall KJ;
```

```
PI Mcarthur HAI;
XX
XX WPI; 2000-182117/16.
XX
XX Mutated Type I polyketide synthase containing a polylinker site in an
PT extension module for replacement of a reductive loop sequence, for
PT producing polyketides, e.g. B1 avermectin.
XX
XX Disclosure; Fig 7a-f; 75pp; English.
XX
XX The present sequence is that of DNA encoding the first 2 modules of the
CC avermectin polyketide synthase (PKS) of Streptomyces avermectilis. The
CC invention relates to nucleic acids encoding a Type I PKS such as
CC avermectin in which a polylinker with multiple restriction sites replaces
CC or 1 more PKS genes encoding enzymes associated with reduction. Novel PKS
CC are provided in which in which the reductive loop in a selected module of
CC the Type I PKS is replaced with the equivalent segment from the same or
CC different PKS gene cluster or by a mutated or synthetic segment. Vectors
CC and host cells, and methods for producing novel polyketides by culturing
CC host cells are claimed. The polyketides obtained are useful as
CC antibiotics and insecticides. Fermentation products containing C22-C23
CC dihydroavermectin, ivermectin and B1 avermectins are claimed
XX
XX Sequence 12381 BP; 1884 A; 4561 C; 4005 G; 1931 T; 0 U; 0 Other;
SQ
Query Match 2.7%; Score 22; DB 3; Length 12381;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 143 AGCGCGCGACCGGCTCACC GA 164
DB 333 AGCGCGCGACCGGCTCACC GA 312
RESULT 11
ABV75561
ID ABV75561 standard; DNA; 20 BP.
XX
XX ABV75561;
AC
XX
XX 22-JAN-2003 (first entry)
DT
XX Saccharopolyspora spnF forward PCR primer.
XX
XX Butenyl; biosynthetic enzyme; PKS; polyketide synthetase; macrolide;
XX metabolite; spinosyn; PCR; primer; ss; spnF.
XX
XX Saccharopolyspora spinosa.
OS
XX WO200279477-A2.
XX
XX 10-OCT-2002.
XX
XX 28-MAR-2002; 2002WO-US009968.
XX
XX 30-MAR-2001; 2001US-0280175P.
XX
XX (DOWC ) DOW AGROSCIENCES LLC.
XX
XX Hahn DR, Jackson JD, Bullard BS, Gustafson GD, Waldron C;
PI Mitchell JC;
XX
XX WPI; 2003-058434/05.
XX
XX New butenyl-spinosyn biosynthetic genes, useful for increasing the
PT production of butenyl-spinosyn insecticidal macrolides, or for changing
PT the metabolites or products produced by spinosyn-producing
PT microorganisms.
XX
XX Example 1; Page 217; 218pp; English.
XX
XX The invention relates to a novel DNA molecule comprising a DNA sequence
CC that encodes a butenyl-spinosyn biosynthetic enzyme, a butenyl-spinosyn
```

CC PKS (polyketide synthetase) domain, or a spinosyn PKS module. The butenyl
CC -spinosyn biosynthetic genes are useful for increasing the production of
CC butenyl-spinosyn insecticidal macrocides. The genes are also useful for
CC changing the metabolites or products produced by spinosyn-producing
CC microorganisms. The present sequence represents a PCR primer used to
CC amplify the spinosyn biosynthetic gene spnf from a *S. spinosa*
CC biosynthetic gene cluster
SQ Sequence 20 BP; 6 A; 7 C; 5 G; 2 T; 0 U; 0 Other;
Query Match 2.4%; Score 20; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 254 GCGACNACGCGATCCGATC 273
|||||
1 GCGACNACGCGATCCGATC 20
DB
RESULT 12
ADP62503
ID ADP62503 standard; cDNA; 237 BP.
XX
AC ADP62503;
XX
DT 09-SEP-2004 (first entry)
XX
DE Maize carbon assimilation pathway enzyme cDNA #2643.
XX
KM Carbon assimilation pathway enzyme; gene; ss; maize; corn;
KM ribulose-bisphosphate carboxylase; phosphoglycerate kinase;
KM glyceraldehyde 3-phosphate dehydrogenase; triose phosphate isomerase;
KM aldolase; fructose-1,6-bisphosphate; transketolase;
KM sedohepuloose-1,7-bisphosphate; D-ribulose-5-phosphate kinase;
KM ribose-5-phosphate isomerase; ribose-5-phosphate kinase;
KM phosphoenolpyruvate carboxylase; NADP-dependent malate dehydrogenase;
KM aspartate aminotransferase; alanine aminotransferase;
KM NADP-dependent malic enzyme; NAD-dependent malic enzyme;
KM PEP carboxykinase; pyruvate; phosphate dikinase; pyrophosphatase;
KM plant metabolic pathway; plant breeding.
XX
OS Zea mays.
XX
PN US2004116682-A1.
XX
PD 17-JUN-2004.
XX
PF 16-NOV-2001; 2001US-00987899.
XX
PR 06-MAR-1998; 98US-0076912P.
PR 04-MAR-1999; 99US-00262979.
XX
PA (CHEI/) CHEIKH N.
PA (MILL/) MILLER P W.
PA (OCON/) O'CONNELL K M.
PA (LIU/) LIU J.
PI Cheikh N, Miller PW, O'Connell KM, Liu J;
PI WPI; 2004-498291/47.
DR
XX
PT New substantially purified nucleic acid molecule encoding a maize or
PT soybean carbon assimilation pathway enzyme, useful for gene mapping,
PT identification and analysis, plant breeding, and preparation of
PT constructs.
XX
PS Claim 2; SEQ ID NO 5572; 196pp; English.
XX
CC The invention relates to a substantially purified nucleic acid molecule
CC that encodes a maize or soybean carbon assimilation pathway enzyme or its
CC fragment. The maize or soybean carbon assimilation pathway enzyme or its
CC fragment is selected from ribulose-bisphosphate carboxylase,
CC phosphoglycerate kinase, glyceraldehyde 3-phosphate dehydrogenase, triose

CC phosphate isomerase, aldolase, fructose-1,6-bisphosphatase,
CC transketolase, sedohepuloose-1,7-bisphosphatase, D-ribulose-5-phosphate-3
CC epimerase, ribose-5-phosphate isomerase, ribose-5-phosphate kinase,
CC phosphoenolpyruvate carboxylase, NADP-dependent malate dehydrogenase,
CC aspartate aminotransferase, alanine aminotransferase, NADP-dependent
CC malic enzyme, NAD-dependent malic enzyme, PEP carboxykinase, pyruvate,
CC phosphate dikinase and pyrophosphatase. The invention also relates to a
CC substantially purified antibody or its fragment which is capable of
CC specifically binding to a specific maize or soybean carbon assimilation
CC pathway enzyme or its fragment, a transformed plant having a nucleic acid
CC molecule comprising an exogenous promoter region which functions in a
CC plant cell to cause the production of an mRNA molecule, and a method of
CC determining a level or pattern in a plant cell of a carbon assimilation
CC pathway enzyme in a plant metabolic pathway. The methods and compositions
CC of the invention are useful for gene mapping, gene identification and
CC analysis, plant breeding and preparation of constructs for use in plant
CC gene expression and transgenic plants. This sequence represents cDNA
CC encoding a maize carbon assimilation pathway enzyme of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 237 BP; 28 A; 101 C; 71 G; 37 T; 0 U; 0 Other;
Query Match 2.4%; Score 20; DB 12; Length 237;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 492 CATCTCGGCGTCACCGAG 511
|||||
192 CATCTCGGCGTCACCGAG 211
DB
RESULT 13
ADP62374
ID ADP62374 standard; cDNA; 263 BP.
XX
AC ADP62374;
XX
DT 09-SEP-2004 (first entry)
XX
DE Maize carbon assimilation pathway enzyme cDNA #2514.
XX
KM Carbon assimilation pathway enzyme; gene; ss; maize; corn;
KM ribulose-bisphosphate carboxylase; phosphoglycerate kinase;
KM glyceraldehyde 3-phosphate dehydrogenase; triose phosphate isomerase;
KM aldolase; fructose-1,6-bisphosphate; transketolase;
KM sedohepuloose-1,7-bisphosphate; D-ribulose-5-phosphate kinase;
KM ribose-5-phosphate isomerase; ribose-5-phosphate kinase;
KM phosphoenolpyruvate carboxylase; NADP-dependent malate dehydrogenase;
KM aspartate aminotransferase; alanine aminotransferase;
KM NADP-dependent malic enzyme; NAD-dependent malic enzyme;
KM PEP carboxykinase; pyruvate; phosphate dikinase; pyrophosphatase;
KM plant metabolic pathway; plant breeding.
XX
OS Zea mays.
XX
PN US2004116682-A1.
XX
PD 17-JUN-2004.
XX
PF 16-NOV-2001; 2001US-00987899.
XX
PR 06-MAR-1998; 98US-0076912P.
PR 04-MAR-1999; 99US-00262979.
XX
PA (CHEI/) CHEIKH N.
PA (MILL/) MILLER P W.
PA (OCON/) O'CONNELL K M.
PA (LIU/) LIU J.
PI Cheikh N, Miller PW, O'Connell KM, Liu J;

DR WPI; 2004-498291/47.
 XX
 XX New substantially purified nucleic acid molecule encoding a maize or
 PT soybean carbon assimilation pathway enzyme, useful for gene mapping, gene
 PT identification and analysis, plant breeding, and preparation of
 PT constructs.
 PS
 PS Claim 2; SEQ ID NO 5443; 196pp; English.
 XX
 XX The invention relates to a substantially purified nucleic acid molecule
 CC that encodes a maize or soybean carbon assimilation pathway enzyme or its
 CC fragment. The maize or soybean carbon assimilation pathway enzyme or its
 CC fragment is selected from ribulose-bisphosphate carboxylase,
 CC phosphoglycerate kinase, glyceraldehyde 3-phosphate dehydrogenase, triose
 CC phosphate isomerase, aldolase, fructose-1,6-bisphosphatase, triose
 CC transketolase, sedohepuloose-1,7-bisphosphatase, D-ribulose-5-phosphate-3-
 CC epimerase, ribose-5-phosphate isomerase, ribose-5-phosphate kinase,
 CC phosphoenolpyruvate carboxylase, NADP-dependent malate dehydrogenase,
 CC aspartate aminotransferase, alanine aminotransferase, NADP-dependent
 CC malic enzyme, NAD-dependent malic enzyme, PEP carboxylase, pyruvate,
 CC phosphate dikinase and pyrophosphatase. The invention also relates to a
 CC substantially purified antibody or its fragment which is capable of
 CC specifically binding to a specific maize or soybean carbon assimilation
 CC pathway enzyme or its fragment, a transformed plant having a nucleic acid
 CC molecule comprising an exogenous promoter region which functions in a
 CC plant cell to cause the production of an mRNA molecule, and a method of
 CC determining a level or pattern in a plant cell of a carbon assimilation
 CC pathway enzyme in a plant metabolic pathway. The methods and compositions
 CC of the invention are useful for gene mapping, gene identification and
 CC analysis, plant breeding and preparation of constructs for use in plant
 CC gene expression and transgenic plants. This sequence represents cDNA
 CC encoding a maize carbon assimilation pathway enzyme of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html.
 CC
 XX
 SQ Sequence 263 BP; 33 A; 105 C; 82 G; 43 T; 0 U; 0 Other;
 Query Match 2.4%; Score 20; DB 12; Length 263;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 492 CATCTCGGCGTCAACGAGG 511
 Db |||||
 162 CATCTCGGCGTCAACGAGG 181
 RESULT 14
 ADP62375 standard; cDNA; 290 BP.
 ID ADP62375;
 AC ADP62375;
 XX
 DT 09-SEP-2004 (first entry)
 XX
 DE Maize carbon assimilation pathway enzyme cDNA #2515.
 XX
 XX Carbon assimilation pathway enzyme; gene; ss; maize; corn;
 KM ribulose-bisphosphate carboxylase; phosphoglycerate kinase;
 KM glyceraldehyde 3-phosphate dehydrogenase; triose phosphate isomerase;
 KM aldolase; fructose-1,6-bisphosphatase; transketolase;
 KM sedohepuloose-1,7-bisphosphatase; D-ribulose-5-phosphate-3-epimerase;
 KM ribose-5-phosphate isomerase; ribose-5-phosphate kinase;
 KM phosphoenolpyruvate carboxylase; NADP-dependent malate dehydrogenase;
 KM aspartate aminotransferase; alanine aminotransferase;
 KM NADP-dependent malic enzyme; NAD-dependent malic enzyme;
 KM PEP carboxylase; pyruvate; phosphate dikinase; pyrophosphatase;
 KM plant metabolic pathway; plant breeding.
 XX
 OS Zea mays.
 XX
 XX US2004116682-A1.
 KM
 XX

PD 17-JUN-2004.
 XX
 XX 16-NOV-2001; 2001US-00987899.
 PF
 XX 06-MAR-1998; 98US-0076912P.
 PR
 PR 04-MAR-1999; 99US-00262979.
 XX
 XX (CHEI/) CHEIKH N.
 PA (MILL/) MILLER P W.
 PA (OCON/) O'CONNELL K M.
 PA (LIU/) LIU J.
 XX
 PI Cheikh N, Miller PW, O'Connell KM, Liu J;
 DR WPI; 2004-498291/47.
 XX
 XX New substantially purified nucleic acid molecule encoding a maize or
 PT soybean carbon assimilation pathway enzyme, useful for gene mapping, gene
 PT identification and analysis, plant breeding, and preparation of
 PT constructs.
 PS
 PS Claim 2; SEQ ID NO 5444; 196pp; English.
 XX
 XX The invention relates to a substantially purified nucleic acid molecule
 CC that encodes a maize or soybean carbon assimilation pathway enzyme or its
 CC fragment. The maize or soybean carbon assimilation pathway enzyme or its
 CC fragment is selected from ribulose-bisphosphate carboxylase,
 CC phosphoglycerate kinase, glyceraldehyde 3-phosphate dehydrogenase, triose
 CC phosphate isomerase, aldolase, fructose-1,6-bisphosphatase, triose
 CC transketolase, sedohepuloose-1,7-bisphosphatase, D-ribulose-5-phosphate-3-
 CC epimerase, ribose-5-phosphate isomerase, ribose-5-phosphate kinase,
 CC phosphoenolpyruvate carboxylase, NADP-dependent malate dehydrogenase,
 CC aspartate aminotransferase, alanine aminotransferase, NADP-dependent
 CC malic enzyme, NAD-dependent malic enzyme, PEP carboxylase, pyruvate,
 CC phosphate dikinase and pyrophosphatase. The invention also relates to a
 CC substantially purified antibody or its fragment which is capable of
 CC specifically binding to a specific maize or soybean carbon assimilation
 CC pathway enzyme or its fragment, a transformed plant having a nucleic acid
 CC molecule comprising an exogenous promoter region which functions in a
 CC plant cell to cause the production of an mRNA molecule, and a method of
 CC determining a level or pattern in a plant cell of a carbon assimilation
 CC pathway enzyme in a plant metabolic pathway. The methods and compositions
 CC of the invention are useful for gene mapping, gene identification and
 CC analysis, plant breeding and preparation of constructs for use in plant
 CC gene expression and transgenic plants. This sequence represents cDNA
 CC encoding a maize carbon assimilation pathway enzyme of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html.
 CC
 XX
 SQ Sequence 290 BP; 38 A; 123 C; 74 G; 55 T; 0 U; 0 Other;
 Query Match 2.4%; Score 20; DB 12; Length 290;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 492 CATCTCGGCGTCAACGAGG 511
 Db |||||
 240 CATCTCGGCGTCAACGAGG 259
 RESULT 15
 ADP62373 standard; cDNA; 294 BP.
 ID ADP62373;
 AC ADP62373;
 XX
 DT 09-SEP-2004 (first entry)
 XX
 DE Maize carbon assimilation pathway enzyme cDNA #2513.
 XX
 XX Carbon assimilation pathway enzyme; gene; ss; maize; corn;
 KM ribulose-bisphosphate carboxylase; phosphoglycerate kinase;
 KM

Search completed: August 29, 2005, 17:33:22
JOD time : 591 secs

KW glyceralddehyde 3-phosphate dehydrogenase; triose phosphate isomerase;
KW aldolase; fructose-1,6-bisphosphate; transketolase;
KW sedohepulo-1,7-bisphosphate; D-ribulose-5-phosphate-3-epimerase;
KW ribose-5-phosphate isomerase; ribose-5-phosphate kinase;
KW phosphoenolpyruvate carboxylase; NADP-dependent malate dehydrogenase;
KW aspartate aminotransferase; alanine aminotransferase;
KW NADP-dependent malic enzyme; NAD-dependent malic enzyme;
KW PEP carboxykinase; pyruvate; phosphate dikinase; pyrophosphatase;
KW plant metabolic pathway; plant breeding.
OS Zea mays.
XX
XX
PN US2004116682-A1.
PD 17-JUN-2004.
XX
XX
PF 16-NOV-2001; 2001US-00987899.
XX
XX
PR 06-MAR-1998; 98US-0076912P.
PR 04-MAR-1999; 99US-00262979.
XX
XX
PA (CHEI/) CHEIKH N.
PA (MILL/) MILLER P W.
PA (OCON/) O'CONNELL K M.
PA (LIU/) LIU J.
XX
XX
PI Cheikh N, Miller PW, O'Connell KM, Liu J;
XX
XX
DR WPI; 2004-498291/47.
XX
XX
PT New substantially purified nucleic acid molecule encoding a maize or
PT soybean carbon assimilation pathway enzyme, useful for gene mapping, gene
PT identification and analysis, plant breeding, and preparation of
PT constructs.
XX
XX
PS Claim 2; SEQ ID NO 5442; 196bp; English.
XX
XX
CC The invention relates to a substantially purified nucleic acid molecule
CC that encodes a maize or soybean carbon assimilation pathway enzyme or its
CC fragment. The maize or soybean carbon assimilation pathway enzyme or its
CC fragment is selected from ribulose-bisphosphate carboxylase,
CC phosphoglycerate kinase, glyceraldehyde 3-phosphate dehydrogenase, triose
CC phosphate isomerase, aldolase, fructose-1,6-bisphosphatase, transketolase,
CC transketolase, sedohepulo-1,7-bisphosphatase, D-ribulose-5-phosphate-3-
CC -epimerase, ribose-5-phosphate isomerase, ribose-5-phosphate kinase,
CC phosphoenolpyruvate carboxylase, NADP-dependent malate dehydrogenase,
CC aspartate aminotransferase, alanine aminotransferase, NADP-dependent
CC malic enzyme, NAD-dependent malic enzyme, PEP carboxykinase, pyruvate,
CC phosphate dikinase and pyrophosphatase. The invention also relates to a
CC substantially purified antibody or its fragment which is capable of
CC specifically binding to a specific maize or soybean carbon assimilation
CC pathway enzyme or its fragment, a transformed plant having a nucleic acid
CC molecule comprising an exogenous promoter region which functions in a
CC plant cell to cause the production of an mRNA molecule, and a method of
CC determining a level or pattern in a plant cell of a carbon assimilation
CC pathway enzyme in a plant metabolic pathway. The methods and compositions
CC of the invention are useful for gene mapping, gene identification and
CC analysis, plant breeding and preparation of constructs for use in plant
CC gene expression and transgenic plants. This sequence represents cDNA
CC encoding a maize carbon assimilation pathway enzyme of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html.
XX
XX
SQ Sequence 294 BP; 36 A; 127 C; 77 G; 54 T; 0 U; 0 Other;

Query Match 2.4%; Score 20; DB 12; Length 294;

Best Local Similarity 100.0%; Pred. No. 27;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 492 CATCTCGGCTCACCGAGG 511
|||
DB 249 CATCTCGGCTCACCGAGG 268

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: August 29, 2005, 17:08:17 ; Search time 3416 Seconds
(without alignments)
9226.351 Million cell updates/sec

Title: US-10-069-353A-7

Sequence: 828 1 gtcgtgcgcagctgcgcacc.....cgcggaagccgctgcgctga 828

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hlc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	21	2.5	243	7	CK916588	CK916588 p3imgcf_0
2	21	2.5	436	7	CK917095	CK917095 p3imgcf_0
3	21	2.5	446	1	AU195670	AU195670 p3imgcf_0
4	21	2.5	495	7	CK916732	CK916732 p3imgcf_0
5	21	2.5	721	1	AU168919	AU168919 p3imgcf_0
6	21	2.5	892	9	CNS02017	AL208708 Tetradon
7	20	2.4	122	2	AM799865	AM799865 PM1-UM005
8	20	2.4	213	5	BQ466968	BQ466968 HS02C10r
9	20	2.4	253	6	C23552	C23552 C23552 R1ce
10	20	2.4	309	1	AJ609797	AJ609797 AJ609797
11	20	2.4	341	1	AU070566	AU070566 AU070566
12	20	2.4	367	9	CC623049	CC623049 OGBRBJ16TV
13	20	2.4	376	6	CD951480	CD951480 SAM_67 Ge
14	20	2.4	413	6	CD961336	CD961336 SDU_109 G
15	20	2.4	453	6	CA501092	CA501092 WHE4029_B
16	20	2.4	465	2	AU181996	AU181996 AU181996
17	20	2.4	465	2	BE403165	BE403165 WHE0426_B
18	20	2.4	466	7	CO519448	CO519448 3530_1_12
19	20	2.4	475	5	BU982086	BU982086 HA25H18r
20	20	2.4	476	2	AM700130	AM700130 GB34H03.Y
21	20	2.4	476	6	CB882644	CB882644 H102H08w
22	20	2.4	508	4	BG366814	BG366814 HVGME1000
23	20	2.4	517	5	BQ466703	BQ466703 HS01F02T
24	20	2.4	562	7	CO523560	CO523560 3530_1_15

25	20	2.4	567	6	CA137413	CA137413 SCAGRT203
26	20	2.4	580	6	CD998957	CD998957 QBF2h02.x
27	20	2.4	580	9	CR101195	CR101195 Reverse s
28	20	2.4	597	6	CA172082	CA172082 SCFSB107
29	20	2.4	597	6	CB883052	CB883052 H001A18w
30	20	2.4	613	6	CA067741	CA067741 SCQSDAD105
31	20	2.4	620	8	BZ716626	BZ716626 OGEAC41TM
32	20	2.4	621	5	BU102550	BU102550 SCCCAD100
33	20	2.4	621	6	CA064657	CA064657 SCCCAD100
34	20	2.4	621	6	CA014854	CA014854 HT12115r
35	20	2.4	629	8	BZ716620	BZ716620 OGEAC41TC
36	20	2.4	639	6	CD998956	CD998956 OBF2h02.p
37	20	2.4	650	6	CA014703	CA014703 HT12B20r
38	20	2.4	670	6	CD878187	CD878187 AZ04_102C
39	20	2.4	672	6	CA163453	CA163453 SCRLR2311
40	20	2.4	682	6	CA066408	CA066408 SCEQAD101
41	20	2.4	693	5	BQ466646	BQ466646 HS01C10T
42	20	2.4	695	6	CD872088	CD872088 AZ02_119M
43	20	2.4	712	6	CA185154	CA185154 SCBSST309
44	20	2.4	712	6	CD440782	CD440782 EL01N0560
45	20	2.4	758	6	CA130606	CA130606 SCCCRT100

ALIGNMENTS

RESULT 1
CK916588
LOCUS
DEFINITION
CK916588
ACCESSION
CK916588
VERSION
CK916588.1
KEYWORDS
SOURCE
ORGANISM
EST.
Magnaporthe grisea (anamorph: Pyricularia grisea)
Magnaporthe grisea
Bukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.

1 (bases 1 to 243)
Dong,H., Peng,Y., Chen,B., Li,Y. and Li,D.
large-scale identification of ESTs from Magnaporthe grisea by
normalized cDNA library sequencing (2004b)
Unpublished (2004)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Debao Li
Bioinformatics and Gene Network Research Group
Zhejiang University, China Agricultural University, Guangxi
University
Kaixuan Road 268#, Hangzhou, Zhejiang, 310029, china
Tel: 0086-571-86961525
Fax: 0086-571-86971183
Email: webmaster@estarray.org, URL: http://www.estarray.org
Seq primer: M13 reverse primer.
Location/Qualifiers

FEATURES
source
1..243
/organism="Magnaporthe grisea"
/mol_type="mRNA"
/db_xref="taxon:148305"
/clone="p3imgcf_003617"
/issue_type="Myceium", conidium, germinating conidium,
swelling appressorium, mature appressorium, penetration
peg"
/dev_stage="Myceium, conidium, germinating conidium,
swelling appressorium, mature appressorium, penetration
peg"
/clone_1ib="Normalized Magnaporthe grisea cDNA pRS423
library"
/note="Vector: pRS423"

ORIGIN

Query Match 2.5%; Score 21; DB 7; Length 243;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 516 CAACGAGAGCGGCGCGG 536
 |||||
 Db 77 CAACGAGAGCGGCGCGG 97

RESULT 2
 LOCUS CK917095 436 bp mRNA linear EST 22-APR-2004
 DEFINITION p3fmgcf_004212 Normalized Magnaporthe grisea cDNA PRS423 library
 MAGNAPORTHE GRISEA cDNA clone p3fmgcf_004212, mRNA sequence.

ACCESSION CK917095
 VERSION CK917095.1 GI:45375800
 KEYWORDS EST.
 SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)
 ORGANISM Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes; incertae sedis; Magnaporthaceae; Magnaporthe.
 1 (bases 1 to 436)
 REFERENCE Dong,H., Peng,Y., Chen,B., Li,Y. and Li,D.
 Large-scale identification of ESTs from Magnaporthe grisea by
 normalized cDNA library sequencing (2004b)
 JOURNAL Unpublished (2004)
 COMMENT Contact: Debao Li
 Bioinformatics and Gene Network Research Group
 Zhejiang University, China Agricultural University, Guangxi
 University
 Kaixuan Road 268#, Hangzhou, Zhejiang, 310029, china
 Tel: 0086-571-86961525
 Fax: 0086-571-86971183
 Email: webmaster@estarray.org, URL: http://www.estarray.org
 Seq primer: M13 reverse primer.
 Location/Qualifiers
 1..436
 /organism="Magnaporthe grisea"
 /mol_type="mRNA"
 /db_xref="taxon:148305"
 /clone="p3fmgcf_004212"
 /tissue_type="mycelium, conidium, germinating conidium,
 swelling appressorium, mature appressorium, penetration
 peg"
 /dev_stage="mycelium, conidium, germinating conidium,
 swelling appressorium, mature appressorium, penetration
 peg"
 /clone_lib="Normalized Magnaporthe grisea cDNA PRS423
 library"
 /note="Vector: pRS423"

ORIGIN
 Query Match 2.5%; Score 21; DB 7; Length 436;
 Best Local Similarity 100.0%; Pred. No. 9.9;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 516 CAACGAGAGCGGCGCGG 536
 |||||
 Db 269 CAACGAGAGCGGCGCGG 289

RESULT 3
 LOCUS AU195670/c 446 bp mRNA linear EST 14-OCT-2003
 DEFINITION AU195670 Porphyra yezoensis TU-1 sporophytes Porphyra yezoensis
 cDNA clone PFL078F08_r_5', mRNA sequence.

ACCESSION AU195670
 VERSION AU195670.1 GI:31937533
 KEYWORDS EST.
 SOURCE Porphyra yezoensis
 ORGANISM Porphyra yezoensis
 Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
 Porphyra.
 1 (bases 1 to 446)
 REFERENCE Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and
 Tabata,S.
 COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS

OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG
 FREQUENCY ANALYSIS
 J. Phycol. 39 (5), 923-930 (2003)
 JOURNAL U. Physiol. 39 (5), 923-930 (2003)
 COMMENT Contact: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yama 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
 source 1..446
 /organism="Porphyra yezoensis"
 /mol_type="mRNA"
 /strain="TU-1"
 /db_xref="taxon:2788"
 /clone="PFL078F08_r"
 /dev_stage="sporophytes"
 /clone_lib="Porphyra yezoensis TU-1 sporophytes"

ORIGIN
 Query Match 2.5%; Score 21; DB 1; Length 446;
 Best Local Similarity 100.0%; Pred. No. 9.9;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 499 GCGCTACCGAGTCTGTCAAA 519
 |||||
 Db 99 GCGCTACCGAGTCTGTCAAA 79

RESULT 4
 LOCUS CK916732 495 bp mRNA linear EST 22-APR-2004
 DEFINITION p3fmgcf_003787 Normalized Magnaporthe grisea cDNA PRS423 library
 MAGNAPORTHE GRISEA cDNA clone p3fmgcf_003787, mRNA sequence.

ACCESSION CK916732
 VERSION CK916732.1 GI:45375437
 KEYWORDS EST.
 SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)
 ORGANISM Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes; incertae sedis; Magnaporthaceae; Magnaporthe.
 1 (bases 1 to 495)
 REFERENCE Dong,H., Peng,Y., Chen,B., Li,Y. and Li,D.
 Large-scale identification of ESTs from Magnaporthe grisea by
 normalized cDNA library sequencing (2004b)
 JOURNAL Unpublished (2004)
 COMMENT Contact: Debao Li
 Bioinformatics and Gene Network Research Group
 Zhejiang University, China Agricultural University, Guangxi
 University
 Kaixuan Road 268#, Hangzhou, Zhejiang, 310029, china
 Tel: 0086-571-86961525
 Fax: 0086-571-86971183
 Email: webmaster@estarray.org, URL: http://www.estarray.org
 Seq primer: M13 reverse primer.
 Location/Qualifiers
 1..495
 /organism="Magnaporthe grisea"
 /mol_type="mRNA"
 /db_xref="taxon:148305"
 /clone="p3fmgcf_003787"
 /tissue_type="mycelium, conidium, germinating conidium,
 swelling appressorium, mature appressorium, penetration
 peg"
 /dev_stage="mycelium, conidium, germinating conidium,
 swelling appressorium, mature appressorium, penetration
 peg"
 /clone_lib="Normalized Magnaporthe grisea cDNA PRS423
 library"
 /note="Vector: pRS423"

ORIGIN
 Query Match 2.5%; Score 21; DB 7; Length 495;
 Best Local Similarity 100.0%; Pred. No. 9.8;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 516 CAACGAGAGCGGGCGGG 536
 |||||
 332 CAACGAGAGCGGGCGGG 352

RESULT 5
 AUI68919/c 721 bp mRNA linear EST 29-JAN-2001
 LOCUS AUI68919 OI-br-ad cDNA Oryzias latipes cDNA clone br3805, mRNA
 DEFINITION sequence.

ACCESSION AUI68919
 VERSION AUI68919.1 GI:12590988
 KEYWORDS EST.
 SOURCE Oryzias latipes (Japanese medaka)
 ORGANISM Oryzias latipes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
 Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
 1 (bases 1 to 721)
 Mita, K., Ishikawa, Y. and Yamauchi, M.
 Establishment of cDNA database of medaka, Oryzias latipes
 Unpublished (2001)
 CONTACT: Mita K
 Genome Research Group
 National Institute of Agrobiological Sciences
 Otsu 1-2, Tsukuba, Ibaraki 305-8634, Japan
 Email: kmika@nias.affrc.go.jp
 method: uni-directional sequence direction: sequenced from T3 primer
 (5' -> 3').

FEATURES
 source
 1. 721
 /organism="Oryzias latipes"
 /mol_type="mRNA"
 /strain="HNT"
 /db_xref="taxon:8090"
 /clone="br3805"
 /sex="female/male mixed"
 /tissue_type="brain"
 /dev_stage="adult"
 /clone_1ib="OI-br-ad cDNA"

ORIGIN
 Query Match 2.5%; Score 21; DB 1; Length 721;
 Best Local Similarity 100.0%; Pred. No. 9 6;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 GGGCTTCTGGCAGCAGCGCG 148
 |||||
 583 GGGCTTCTGGCAGCAGCGCG 563

RESULT 6
 CNGS02Q17 892 bp DNA linear GSS 01-SEP-2000
 LOCUS CNGS02Q17 Tetraodon nigroviridis genome survey sequence PUC-Or1 end of clone
 DEFINITION 156N1 of library G from Tetraodon nigroviridis, genomic survey
 sequence.
 ACCESSION AL208708
 VERSION AL208708.1 GI:7867527
 KEYWORDS GSS; genome survey sequence.
 SOURCE Tetraodon nigroviridis
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodontidae; Tetraodon.
 1
 Roest Crolius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,
 Bernot, A., Fizames, C., Wincker, P., Brothier, P., Quetier, F.,
 Saurin, W. and Weissenbach, J.

TITLE Estimate of human gene number provided by genome-wide analysis
 using Tetraodon nigroviridis DNA sequence
 JOURNAL Nat. Genet. 25 (2), 235-238 (2000)
 MEDLINE 20296633
 PUBMED 10835645
 AUTHORS Roest Crolius, H., Jallion, O., Dasilva, C., Ozouf-Costaz, C.,
 Fizames, C., Bernot, A. and Weissenbach, J.
 Saurin, W., Bernot, A. and Weissenbach, J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 Genome Res. 10 (7), 939-949 (2000)

JOURNAL 20359837
 MEDLINE 10892143
 PUBMED 3 (bases 1 to 892)
 REFERENCE Genoscope.
 AUTHORS Direct Submission
 TITLE Submitted (12-APR-2000) Genoscope - Centre National de Séquençage :
 BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetraodon.
 location/Qualifiers
 1. 892
 /organism="Tetraodon nigroviridis"
 /mol_type="genomic DNA"
 /db_xref="taxon:99883"
 /clone="156N10"
 /clone_1ib="G"
 /note="TGenoscope sequence ID : C0AG156DG05SP1-end :
 PUC-Or1"

ORIGIN
 Query Match 2.5%; Score 21; DB 9; Length 892;
 Best Local Similarity 100.0%; Pred. No. 9 5;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 543 GGTGTCCGGGAGCAGGTGCC 563
 |||||
 653 GGTGTCCGGGAGCAGGTGCC 673

RESULT 7
 AW799865 122 bp mRNA linear EST 16-MAY-2000
 LOCUS PML-UM0056-100300-003-f06 UM0056 Homo sapiens cDNA, mRNA sequence.
 DEFINITION AW799865
 ACCESSION AW799865.1 GI:7851735
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 122)
 Dias Neto, E., Garcia Correia, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Branstetter, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.U.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL 20202663
 MEDLINE 10737800
 PUBMED
 COMMENT Contact: Simpson A.U.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922

Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=et2=PM1-UM0056-100
 300-003-f06&cl=2000-03-10&cl=1)
 Seq primer: puc 18 forward
 High quality sequence start: 7
 High quality sequence stop: 71.
 Location/Qualifiers

FEATURES

source

```

1..122
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="UM0056"
/note="Organ: uterus; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

```

ORIGIN

Query Match 2.4%; Score 20; DB 2; Length 122;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 604 TCCTGCGGTGACGCGGGTT 623
 ||||||||||||||||
 Db 99 TCGCTGCGGTGACGCGGGTT 118

RESULT 8

BQ466968

LOCUS BQ466968 213 bp mRNA linear EST 30-MAY-2002
 DEFINITION HS02C10r HS Hordeum vulgare subsp. vulgare cDNA clone HS02C10

5-PRIME, mRNA sequence.
 BQ466968

ACCESSION BQ466968.1 GI:21274750
 EST.

KEYWORDS

SOURCE

ORGANISM

Hordeum vulgare subsp. vulgare
 Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooideae; Triticeae; Hordeum.
 1 (bases 1 to 213)

REFERENCE

AUTHORS

Zhang, H., Potokina, E., Michalek, W., Weschke, W., Stein, N. and Graner, A.
 Barley ESTs from germinating seeds
 Unpublished (2002)

TITLE

JOURNAL

Contact: Stein Nils
 Molecular Markers Group, Department Genbank
 Institute of Plant Genetics and Crop Plant Research (IPK)
 Corrensstr. 3, 06466, Gatersleben, Germany
 Tel: 039482-5522
 Fax: 039482-5595

Email: stein@ipk-gatersleben.de
 Insert Length: 213 Std Error: 0.00
 Plate: 2 row: C column: 10
 Seq primer: M3rev.

COMMENT

Location/Qualifiers
 1..213
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="barke"
 /sub_species="vulgare"
 /db_xref="taxon:112509"
 /clone="HS02C10"
 /tissue_type="embryo + scutellum"
 /dev_stage="0-16 hours after imbibition"
 /lab_host="X110-Gold"

FEATURES

source

/clone_lib="HS"
 /note="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of cDNA); Site 2: XhoI (3'-end of cDNA); Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also due to the cloning system used Blue/White selection for recombinants is not 100% reliable."

ORIGIN

Query Match 2.4%; Score 20; DB 5; Length 213;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 492 CATCCTCGGCGTCACCGAGG 511
 ||||||||||||||||
 Db 168 CATCCTCGGCGTCACCGAGG 187

RESULT 9

C23552

LOCUS C23552 253 bp mRNA linear EST 03-APR-2002
 DEFINITION C23552 Rice callus Oryza sativa (japonica cultivar-group) cDNA clone C60213_1A, mRNA sequence.

ACCESSION

C23552.1 GI:1944230
 EST.

VERSION

KEYWORDS

SOURCE

ORGANISM

Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartioideae; Oryzaceae; Oryza.
 1 (bases 1 to 253)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Takuji Sasaki
 National Institute of Agrobiological Resources
 Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
 305-8602, Japan
 Tel: 81-288-38-7441
 Fax: 81-288-38-7468
 Email: tsasaki@db.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/PROJECT/'RGP'

FEATURES

source

Location/Qualifiers
 1..253
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="C60213_1A"
 /clone_lib="Rice callus"
 /note="Vector: pBluescript II SK+; Site 1: SalI; Site 2: NotI; cDNA prepared from rice callus mRNAs by using oligo(dT) as a primer and ligating to the SalI-NotI site of pBluescript II SK+ phagemid."

ORIGIN

Query Match

Best Local Similarity

Matches

20; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

OY 492

CATCCTCGGCGTCACCGAGG

511

|||||

Db 147

CATCCTCGGCGTCACCGAGG

166

|||||

RESULT 10

AUF09797

LOCUS AUF09797

DEFINITION AUF09797

Triticum turgidum subsp. durum etiolated seedling 20 day

Triticum turgidum subsp. durum cDNA clone 00262R, mRNA sequence.

309 bp

mRNA

linear

EST 11-DEC-2003

ACCESSION AJ609797
 VERSION AJ609797.1 GI:39726724
 KEYWORDS EST.
 SOURCE Triticum turgidum subsp. durum (durum wheat)
 ORGANISM Triticum turgidum subsp. durum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.
 1 (bases 1 to 309)
 Clifarelli, R.A., D'Onofrio, O. and Lauria, G.
 Expressed Sequences Tags (ESTs) library from totipotent cDNA of
 durum wheat
 JOURNAL Unpublished (2003)
 COMMENT Contact: Clifarelli RA
 Biotechnology
 Metapontum Agrobios
 S.S. Jonica 106 Km 448.2, 75010 Metaponto (MT), Italy.
 Location/Qualifiers
 FEATURES
 source
 1..309
 /organism="Triticum turgidum subsp. durum"
 /mol_type="mRNA"
 /cultivar="Ofanto"
 /sub_species="durum"
 /db_xref="taxon:4567"
 /clone="00262R"
 /issue_type="etiolated seedling"
 /dev_stage="20 day"
 /clone_lib="Triticum turgidum subsp. durum etiolated
 seedling 20 day"
 ORIGIN
 Query Match 2.4%; Score 20; DB 1; Length 309;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 492 CATCTCGGCGTCACCGAG 511
 |||||
 DB 200 CATCTCGGCGTCACCGAG 219
 RESULT 11
 LOCUS AU070566 341 bp mRNA linear EST 02-APR-2002
 DEFINITION AU070566 Rice shoot Oryza sativa (japonica cultivar-group) cDNA
 clone S4279_8A, mRNA sequence.
 ACCESSION AU070566
 VERSION AU070566.1 GI:5005404
 KEYWORDS EST.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 341)
 Yamamoto, K. and Sasaki, T.
 Rice cDNA from etiolated shoot
 Unpublished (1997)
 Contact: Takuji Sasaki
 National Institute of Agrobiological Resources
 Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
 305-8602, Japan
 Tel: 81-298-38-7441
 Fax: 81-298-38-7468
 Email: tsasaki@abrr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/.
 location/Qualifiers
 FEATURES
 source
 1..341
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="S4279_8A"
 /dev_stage="Etiolated shoot (8 days old)"
 /clone_lib="Rice shoot"

ORIGIN /note="Etiolated shoot (8 days old)"
 Query Match 2.4%; Score 20; DB 1; Length 341;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 492 CATCTCGGCGTCACCGAG 511
 |||||
 DB 107 CATCTCGGCGTCACCGAG 126
 RESULT 12
 LOCUS CC623049 367 bp DNA linear GSS 19-JUN-2003
 DEFINITION OGRB12TV ZM 0.7-1.5_KB Zea mays genomic clone ZMMBMA0380D08,
 genomic survey sequence.
 ACCESSION CC623049
 VERSION CC623049.1 GI:31989470
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 367)
 Whitelew, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
 Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, D.A., Rohlfing, T.,
 Citek, R.W., Nuneberg, A., Robbins, D. and Lakey, N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Contact: Cathy Whitelew
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelew@cigr.org
 Seq primer: TF
 Classes: sheared ends.
 Location/Qualifiers
 FEATURES
 source
 1..367
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBMA0380D08"
 /clone_lib="ZM 0.7-1.5_KB"
 /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"
 ORIGIN
 Query Match 2.4%; Score 20; DB 9; Length 367;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 363 CGTCATGCCATGTCCCTGC 382
 |||||
 DB 148 CGTCATGCCATGTCCCTGC 129
 RESULT 13
 LOCUS CD951480 376 bp mRNA linear EST 15-JUL-2003
 DEFINITION SAW 67 GeneTag2 Zea mays cDNA, mRNA sequence.
 ACCESSION CD951480
 VERSION CD951480.1 GI:32799244
 KEYWORDS EST.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 376)

AUTHORS Genoplante.
TITLE Genoplante, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french plant genomics programme 'Genoplante' (<http://www.genoplante.com> and <http://genoplante-info.infobiogen.fr>).
Location/Qualifiers
1..376
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="mixture"
/db_xref="taxon:4577"
/clone_lib="GenetAg2"

ORIGIN
Query Match 2.4%; Score 20; DB 6; Length 376;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 363 CGTCATGCCATGTCCTGC 382
|||||
215 CGTCATGCCATGTCCTGC 234

Db 215 CGTCATGCCATGTCCTGC 234

RESULT 14
LOCUS CD961336 376 bp mRNA linear EST 15-JUL-2003
DEFINITION SDU_109 GenetAg2 Zea mays CDNA, mRNA sequence.
ACCESSION CD961336
VERSION CD961336.1 GI:32809102
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 376)
Genoplante.
TITLE Genoplante, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french plant genomics programme 'Genoplante' (<http://www.genoplante.com> and <http://genoplante-info.infobiogen.fr>).
Location/Qualifiers
1..376
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="mixture"
/db_xref="taxon:4577"
/clone_lib="GenetAg2"

ORIGIN
Query Match 2.4%; Score 20; DB 6; Length 376;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 363 CGTCATGCCATGTCCTGC 382
|||||
215 CGTCATGCCATGTCCTGC 234

Db 215 CGTCATGCCATGTCCTGC 234

RESULT 15
CA501092

LOCUS CA501092 413 bp mRNA linear EST 14-NOV-2002
DEFINITION WHE4029 B12_C23T Wheat meiotic anther cDNA library Triticum aestivum CDNA clone WHE4029_B12_C23, mRNA sequence.
ACCESSION CA501092
VERSION CA501092.1 GI:24992052
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Triticum.
1 (bases 1 to 413)
Anderson, O.D., Chao, S., Crossman, C., Langridge, P., Iazo, G.R., Pham, V., Rausch, C.J., Sutton, T., Woo, J. and Wilson, C.
The structure and function of the expressed portion of the wheat genomes - Meiotic anther cDNA library
Unpublished (2002)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105959773
Fax: 5105959818
Email: oanderson@nps.usda.gov
Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Seq primer: T7 primer.
Location/Qualifiers
1..413
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE4029_B12_C23"
/tissue_type="Anther"
/dex_stage="Meiotic stages pre-meiosis-metaphase I"
/lab_host="E. coli DH10B"
/clone_lib="Wheat meiotic anther cDNA library"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; Plants were grown in a glasshouse. Anther meiotic stage was determined by removing anthers from individual primary florets. One anther was sacrificed for microscopic staging, and if determined to be between (and including) meiotic stages pre-meiosis and metaphase I, the remaining two anthers were collected and pooled for library construction. The tissue, total RNA, and poly(A) RNA were prepared, cDNA synthesised, and directionally ligated into pSPORT1 by Tim Sutton in the P Langridge lab at the Department of Plant Science, University of Adelaide, Waite Campus, Australia. Average insert size 1.5kb. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

ORIGIN
Query Match 2.4%; Score 20; DB 6; Length 413;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 492 CATCTCGGCGTCACGAGG 511
|||||
182 CATCTCGGCGTCACGAGG 201

Db 182 CATCTCGGCGTCACGAGG 201

Search completed: August 29, 2005, 19:35:30
Job time : 3428 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: August 29, 2005, 17:11:57 ; Search time 219 Seconds
(without alignments)
6186.471 Million cell updates/sec

Title: US-10-069-353A-7

Perfect score: 828
1 ggcgtgcgcagctgcgcacc.....cgcggaagccgctgcgctga 828

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 81813359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database :

- Issued Patents NA:*
- 1: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
 - 2: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
 - 3: /cgn2_6/prodata/1/ina/6A.COMB.seq:*
 - 4: /cgn2_6/prodata/1/ina/6B.COMB.seq:*
 - 5: /cgn2_6/prodata/1/ina/PCUS.COMB.seq:*
 - 6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	828	100.0	80161	3 US-09-036-987A-1	Sequence 1, Appl1
2	828	100.0	80161	3 US-09-370-700-1	Sequence 1, Appl1
3	828	100.0	80161	4 US-09-603-207-1	Sequence 1, Appl1
4	20	2.4	3107	2 US-08-813-940-3	Sequence 3, Appl1
5	20	2.4	8051	2 US-08-576-626A-2	Sequence 2, Appl1
6	20	2.4	4403765	3 US-09-103-840A-2	Sequence 2, Appl1
7	20	2.4	4411529	3 US-09-103-840A-1	Sequence 1, Appl1
8	19	2.3	1230025	4 US-09-198-452A-1	Sequence 1, Appl1
9	19	2.3	1230230	4 US-09-438-185A-1	Sequence 1, Appl1
10	18	2.2	266	4 US-09-313-294A-3098	Sequence 3098, Ap
11	18	2.2	272	4 US-09-313-294A-2444	Sequence 2444, Ap
12	18	2.2	282	4 US-09-621-976-489	Sequence 489, Ap
13	18	2.2	391	4 US-09-513-999C-10082	Sequence 10082, A
14	18	2.2	429	4 US-09-252-991A-1854	Sequence 1854, Ap
15	18	2.2	513	4 US-09-252-991A-1154	Sequence 1154, Ap
16	18	2.2	636	4 US-09-170-984-6	Sequence 6, Appl1
17	18	2.2	897	4 US-09-252-991A-9787	Sequence 9787, Ap
18	18	2.2	939	4 US-09-252-991A-10191	Sequence 10191, A
19	18	2.2	945	4 US-09-252-991A-2181	Sequence 2181, Ap
20	18	2.2	1194	4 US-09-252-991A-1785	Sequence 1785, Ap
21	18	2.2	1203	4 US-09-252-991A-2018	Sequence 2018, Ap
22	18	2.2	1386	4 US-09-602-787A-359	Sequence 359, Ap
23	18	2.2	1536	4 US-09-180-109A-3	Sequence 3, Appl1
24	18	2.2	1818	4 US-09-252-991A-1706	Sequence 1706, Ap
25	18	2.2	2052	4 US-09-252-991A-10050	Sequence 10050, A
26	18	2.2	2568	4 US-09-489-039A-2980	Sequence 2980, Ap
27	18	2.2	3981	4 US-09-252-991A-1361	Sequence 1361, Ap

C 28	18	2.2	4179	4 US-09-252-991A-1244	Sequence 1244, Ap
C 29	17	2.1	204	4 US-09-252-991A-11819	Sequence 11819, A
C 30	17	2.1	256	4 US-09-270-767-7360	Sequence 7360, Ap
C 31	17	2.1	256	4 US-09-270-767-22642	Sequence 22642, A
C 32	17	2.1	428	4 US-09-513-999C-2346	Sequence 2346, Ap
C 33	17	2.1	482	4 US-09-621-976-3490	Sequence 3490, Ap
C 34	17	2.1	501	4 US-09-902-540-3908	Sequence 3908, Ap
C 35	17	2.1	531	4 US-09-252-991A-7115	Sequence 7115, Ap
C 36	17	2.1	537	4 US-09-252-991A-5582	Sequence 5582, Ap
37	17	2.1	573	4 US-09-902-540-7223	Sequence 7223, Ap
38	17	2.1	580	1 US-08-078-176-1	Sequence 1, Appl1
39	17	2.1	684	4 US-09-252-991A-2861	Sequence 2861, Ap
C 40	17	2.1	717	4 US-09-902-540-7027	Sequence 7027, Ap
C 41	17	2.1	792	4 US-09-252-991A-7444	Sequence 7444, Ap
42	17	2.1	852	4 US-09-266-965-36	Sequence 36, Appl1
43	17	2.1	888	4 US-09-252-991A-7041	Sequence 7041, Ap
44	17	2.1	927	4 US-09-252-991A-1851	Sequence 1851, Ap
45	17	2.1	960	4 US-09-252-991A-7256	Sequence 7256, Ap

ALIGNMENTS

RESULT 1
US-09-036-987A-1
; Sequence 1, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Brughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; TITLE OF INVENTION: BioSynthetic Genes For Spinosyn Insecticide
; TITLE OF INVENTION: Production
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow Agrosciences LLC Patent Department
; STREET: 9330 Zionville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Releasee #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50, 608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816
; TELEFAX: (317)337-4847
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80161 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-036-987A-1
Query Match 100.0%; Score 828; DB 3; Length 80161;
Best Local Similarity 100.0%; Pred. No. 0;

```
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGTGCAAGTGGCGCACCAACATGCGAGGTTGGGCAATGATGACCGTGTGACG 60
DB 20168 GTGTGCAAGTGGCGCACCAACATGCGAGGTTGGGCAATGATGACCGTGTGACG 20227
QY 61 CCGTTGCTGAATCGGTGCGCGGGCGGCGCCCTGCGCCATCCACCGGCTACTGGAAGAC 120
DB 20228 CCGTTGCTGAATCGGTGCGCGGGCGGCGCCCTGCGCCATCCACCGGCTACTGGAAGAC 20287
QY 121 GACGGGCGGGCTTCCCTGCGAGACAGGCGCGCCGACCCGCTCAACGACCTTGTGCGCAACGG 180
DB 20288 GACGGGCGGGCTTCCCTGCGAGACAGGCGCGCCGACCCGCTCAACGACCTTGTGCGCAACGG 20347
QY 181 ACCGTCCTGATGAGCGGCGCTGCACTGCTGATGTGGGGTGGCGGTACCGGACCAACGAGCG 240
DB 20348 ACCGTCCTGATGAGCGGCGCTGCACTGCTGATGTGGGGTGGCGGTACCGGACCAACGAGCG 20407
QY 241 CTGGCGGTGCGCGCGACCAACCGGATCCAGATCCAGGAGATCCGTCAGCCAGGTGCA 300
DB 20408 CTGGCGGTGCGCGCGACCAACCGGATCCAGATCCAGGAGATCCGTCAGCCAGGTGCA 20467
QY 301 GTGGCCATGCGCGCTGATGTTGGCGGACCGGAACTAAGCACTGAGTGGACTTTCTG 360
DB 20468 GTGGCCATGCGCGCTGATGTTGGCGGACCGGAACTAAGCACTGAGTGGACTTTCTG 20527
QY 361 TCGTCGATGACCATGTCCTGCGCATCCGACCAATGCTTTGACGCGCGCTGGGCGCATG 420
DB 20528 TCGTCGATGACCATGTCCTGCGCATCCGACCAATGCTTTGACGCGCGCTGGGCGCATG 20587
QY 421 CAGTCGCTGTGAGATGTCCGAACCGGACCGGTCCATCCGGGAAATCTTTGAGATCTC 480
DB 20588 CAGTCGCTGTGAGATGTCCGAACCGGACCGGTCCATCCGGGAAATCTTTGAGATCTC 20647
QY 481 AAACCCGGTGGCATCTCTGCGCGCTGACCCGAGTGTCTAAACGAGAACGGGCGCGGAGTG 540
DB 20648 AAACCCGGTGGCATCTCTGCGCGCTGACCCGAGTGTCTAAACGAGAACGGGCGCGGAGTG 20707
QY 541 CCGGTGTCCGGGGAACAGTGGCGGACCGGCTTCCGATCTGCTGGCTGAGCACTTCTG 600
DB 20708 CCGGTGTCCGGGGAACAGTGGCGGACCGGCTTCCGATCTGCTGGCTGAGCACTTCTG 20767
QY 601 GAATGCTGCTGCGACGCGGGGTTCCAGATCTCTGATTTGGAGAGACGTTGTGAGAGAC 660
DB 20768 GAATGCTGCTGCGACGCGGGGTTCCAGATCTCTGATTTGGAGAGACGTTGTGAGAGAC 20827
QY 661 CGGTACTTATGATCCCGAAGTTCCCGAAGAGCTCGCTGCGCACACAGACGGGATCCGAGAC 720
DB 20828 CGGTACTTATGATCCCGAAGTTCCCGAAGAGCTCGCTGCGCACACAGACGGGATCCGAGAC 20887
QY 721 AGGTACGGGCGGCTGTGCGCGGCTGGGCGCGCGGCTGCGATTTATGAAATATATGCC 780
DB 20888 AGGTACGGGCGGCTGTGCGCGGCTGGGCGCGCGGCTGCGATTTATGAAATATATGCC 20947
QY 781 CACGACATGGGCTATGCGATTCTGACGCGCGGAAAGCCGGTGGCTGA 828
DB 20948 CACGACATGGGCTATGCGATTCTGACGCGCGGAAAGCCGGTGGCTGA 20995

RESULT 2
US-09-370-700-1
; Sequence 1, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
```

```
; CURRENT APPLICATION NUMBER: US/09/370,700
; FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/36987
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 80161
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Saccharopolyspora spinosa
US-09-370-700-1

Query Match      100.0%; Score 828; DB 3; Length 80161;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGTGCAAGTGGCGCACCAACATGCGAGGTTGGGCAATGATGACCGTGTGACG 60
DB 20168 GTGTGCAAGTGGCGCACCAACATGCGAGGTTGGGCAATGATGACCGTGTGACG 20227
QY 61 CCGTTGCTGAATCGGTGCGCGGGCGGCGCCCTGCGCCATCCACCGGCTACTGGAAGAC 120
DB 20228 CCGTTGCTGAATCGGTGCGCGGGCGGCGCCCTGCGCCATCCACCGGCTACTGGAAGAC 20287
QY 121 GACGGGCGGGCTTCCCTGCGAGACAGGCGCGCCGACCCGCTCAACGACCTTGTGCGCAACGG 180
DB 20288 GACGGGCGGGCTTCCCTGCGAGACAGGCGCGCCGACCCGCTCAACGACCTTGTGCGCAACGG 20347
QY 181 ACCGTCCTGATGAGCGGCGCTGCACTGCTGATGTGGGGTGGCGGTACCGGACCAACGAGCG 240
DB 20348 ACCGTCCTGATGAGCGGCGCTGCACTGCTGATGTGGGGTGGCGGTACCGGACCAACGAGCG 20407
QY 241 CTGGCGGTGCGCGCGACCAACCGGATCCAGATCCAGGAGATCCGTCAGCCAGGTGCA 300
DB 20408 CTGGCGGTGCGCGCGACCAACCGGATCCAGATCCAGGAGATCCGTCAGCCAGGTGCA 20467
QY 301 GTGGCCATGCGCGCTGATGTTGGCGGACCGGAACTAAGCACTGAGTGGACTTTCTG 360
DB 20468 GTGGCCATGCGCGCTGATGTTGGCGGACCGGAACTAAGCACTGAGTGGACTTTCTG 20527
QY 361 TCGTCGATGACCATGTCCTGCGCATCCGACCAATGCTTTGACGCGCGCTGGGCGCATG 420
DB 20528 TCGTCGATGACCATGTCCTGCGCATCCGACCAATGCTTTGACGCGCGCTGGGCGCATG 20587
QY 421 CAGTCGCTGTGAGATGTCCGAACCGGACCGGTCCATCCGGGAAATCTTTGAGATCTC 480
DB 20588 CAGTCGCTGTGAGATGTCCGAACCGGACCGGTCCATCCGGGAAATCTTTGAGATCTC 20647
QY 481 AAACCCGGTGGCATCTCTGCGCGCTGACCCGAGTGTCTAAACGAGAACGGGCGCGGAGTG 540
DB 20648 AAACCCGGTGGCATCTCTGCGCGCTGACCCGAGTGTCTAAACGAGAACGGGCGCGGAGTG 20707
QY 541 CCGGTGTCCGGGGAACAGTGGCGGACCGGCTTCCGATCTGCTGGCTGAGCACTTCTG 600
DB 20708 CCGGTGTCCGGGGAACAGTGGCGGACCGGCTTCCGATCTGCTGGCTGAGCACTTCTG 20767
QY 601 GAATGCTGCTGCGACGCGGGGTTCCAGATCTCTGATTTGGAGAGACGTTGTGAGAGAC 660
DB 20768 GAATGCTGCTGCGACGCGGGGTTCCAGATCTCTGATTTGGAGAGACGTTGTGAGAGAC 20827
QY 661 CGGTACTTATGATCCCGAAGTTCCCGAAGAGCTCGCTGCGCACACAGACGGGATCCGAGAC 720
DB 20828 CGGTACTTATGATCCCGAAGTTCCCGAAGAGCTCGCTGCGCACACAGACGGGATCCGAGAC 20887
QY 721 AGGTACGGGCGGCTGTGCGCGGCTGGGCGCGCGGCTGCGATTTATGAAATATATGCC 780
DB 20888 AGGTACGGGCGGCTGTGCGCGGCTGGGCGCGCGGCTGCGATTTATGAAATATATGCC 20947
QY 781 CACGACATGGGCTATGCGATTCTGACGCGCGGAAAGCCGGTGGCTGA 828
DB 20948 CACGACATGGGCTATGCGATTCTGACGCGCGGAAAGCCGGTGGCTGA 20995
```



```
RESULT 3
US-09-603-207-1
; Sequence 1, Application US/09603207B
; Patent No. 6521406
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 Div1
; CURRENT APPLICATION NUMBER: US/09/603.207B
; EARLIER FILING DATE: 2000-06-23
; EARLIER APPLICATION NUMBER: 09/370,700
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1
; LENGTH: 80161
; TYPE: DNA
; ORGANISM: Saccharopolyspora spinosa
US-09-603-207-1

Query Match      100.0%; Score 828; DB 4; Length 80161;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GTGTTCGACAGTGGCGGACCAACACTCCACAGCTTGGGAGATGTATGACTGGTACG 60
DB      20168 GTGTTCGACAGTGGCGGACCAACACTCCACAGCTTGGGAGATGTATGACTGGTACG 20227

QY      61 CCGTTGCTGAACCTCGATCGCGGGGCGCCCTGCGCCATCCACACAGGCTACTGGAGAAC 120
DB      20228 CCGTTGCTGAACCTCGATCGCGGGGCGCCCTGCGCCATCCACACAGGCTACTGGAGAAC 20287

QY      121 GACGGGGGCGGCTTCTTGACAGCAGCGCCGACCGGCTCAACGACTTGTGCGCGAAACG 180
DB      20288 GACGGGGGCGGCTTCTTGACAGCAGCGCCGACCGGCTCAACGACTTGTGCGCGAAACG 20347

QY      181 ACCGTGCTCGATGGCGGCGCTTCCACTGCTGATGTGGGGTGGCGGTACCGACACACGCG 240
DB      20348 ACCGTGCTCGATGGCGGCGCTTCCACTGCTGATGTGGGGTGGCGGTACCGACACACGCG 20407

QY      241 CTGGCGGTGGGGGCGGCAACAGCGATCCAGATCAACGCGCACTACCGTCAAGCAGGTCAA 300
DB      20408 CTGGCGGTGGGGGCGGCAACAGCGATCCAGATCAACGCGCACTACCGTCAAGCAGGTCAA 20467

QY      301 GTGGCCATTCGCGCTGATTGGCGACGCGGAAACGCGGACTTAAGCAACGCGGTGGAATTCTG 360
DB      20468 GTGGCCATTCGCGCTGATTGGCGACGCGGAAACGCGGACTTAAGCAACGCGGTGGAATTCTG 20527

QY      361 TGGCTGATGATGATGCTCCCTCCCTTACCCCGACAACTGTTTTCAGCGCGCTGGGCCATG 420
DB      20528 TGGCTGATGATGATGCTCCCTCCCTTACCCCGACAACTGTTTTCAGCGCGCTGGGCCATG 20587

QY      421 CAGTGGCTGTTGAGAGATGTCGGAACCGGACCGTGCATTCGCGGAAATCTTTCAGATCTC 480
DB      20588 CAGTGGCTGTTGAGAGATGTCGGAACCGGACCGTGCATTCGCGGAAATCTTTCAGATCTC 20647

QY      481 AAACCCGGTGGCATCTTCGGGCTCAACGAGGTGCTCAACGAGAAAGCGGCGCGGATG 540
DB      20648 AAACCCGGTGGCATCTTCGGGCTCAACGAGGTGCTCAACGAGAAAGCGGCGCGGATG 20707

QY      541 CCGGTGTCGGGGGACAGTGGCGGACCGGCTTGGATGCTGCGCTGAGCAACTTCTG 600
DB      20708 CCGGTGTCGGGGGACAGTGGCGGACCGGCTTGGATGCTGCGCTGAGCAACTTCTG 20767

QY      601 GAATCGCTGCGTGAAGCGGGGTTGAGATCTTCGATTGGAGAGACGTGTCTCGAGAGAC 660
DB      20768 GAATCGCTGCGTGAAGCGGGGTTGAGATCTTCGATTGGAGAGACGTGTCTCGAGAGAC 20827
```

```
QY      661 CCGTACTTCATGCGCGAGTTCGCGGAGAGACTCGTGCACACACAGCGGATCCGCGAC 720
DB      20828 CCGTACTTCATGCGCGAGTTCGCGGAGAGACTCGTGCACACACAGCGGATCCGCGAC 20887

QY      721 AGGTACGGGCGGCTGTGCGCGGCTGGGCGCGGCTTGGGATTTAGAAATATGCG 780
DB      20888 AGGTACGGGCGGCTGTGCGCGGCTGGGCGCGGCTTGGGATTTAGAAATATGCG 20947

QY      781 CACGACATGGGCTATGCGATTCTGACGGGCGGGAAGCCGCTGGCTGA 828
DB      20948 CACGACATGGGCTATGCGATTCTGACGGGCGGGAAGCCGCTGGCTGA 20995

RESULT 4
US-08-813-940-3/c
; Sequence 3, Application US/08813940
; Patent No. 5834279
; GENERAL INFORMATION:
; APPLICANT: Rubin, Harvey
; APPLICANT: Yang, Jude
; APPLICANT: Avarbock, David
; APPLICANT: Curran, Sean
; TITLE OF INVENTION: Methods of Identifying Compounds that
; TITLE OF INVENTION: Inhibit DNA Synthesis in Mycobacterium Tuberculosis and
; TITLE OF INVENTION: Composition, Reagents and Kits for Performing the Same
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5834279ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/813,940
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,271
; FILING DATE: 07-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: UPAP-0220
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3107 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 207..2909
; US-08-813-940-3

Query Match      2.4%; Score 20; DB 2; Length 3107;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      353 ACTTCTGTCGCTGATGCC 372
DB      2143 ACTTCTGTCGCTGATGCC 2124
```



```
RESULT 5
US-08-576-626A-2
; Sequence 2, Application US/08576626A
; Patent No. 5998194
; GENERAL INFORMATION:
; APPLICANT: Summers, R.G.
; APPLICANT: Katz, L.
; APPLICANT: Donadio, S.
; APPLICANT: Steaver, M.J.
; TITLE OF INVENTION: POLYKETIDE-ASSOCIATED SUGAR
; TITLE OF INVENTION: BIOSYNTHESIS GENES
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/576,626A
; FILING DATE: 21-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Diane Casuto
; REGISTRATION NUMBER: P-40,943
; REFERENCE/DOCKET NUMBER: 5857.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (847) 938-1137
; TELEFAX: (847) 938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8051 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-576-626A-2
Query Match      2.4%; Score 20; DB 2; Length 8051;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      148 GCGACCGGCTCAGCAGCT 167
DB      6799 GCCGACCGGCTCAGCAGCT 6818

RESULT 6
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
US-09-103-840A-2
Query Match      2.4%; Score 20; DB 3; Length 4403765;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      353 ACTTCTCGTGGCTGATGCC 372
DB      4077513 ACTTCTCGTGGCTGATGCC 4077532

RESULT 7
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1
Query Match      2.4%; Score 20; DB 3; Length 4411529;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      353 ACTTCTCGTGGCTGATGCC 372
DB      4085353 ACTTCTCGTGGCTGATGCC 4085352

RESULT 8
US-09-198-452A-1
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffla, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevent
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(15000)
; OTHER INFORMATION: n=a or c or g or t
```



```
; Sequence 489, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTE and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 489
; LENGTH: 282
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..282
; NAME/KEY: misc_feature
; LOCATION: 74
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-489

Query Match      2.2%; Score 18; DB 4; Length 282;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      426 GCGTGTGAGATGTCGA 443
DB      81 GCTGTGGAGATGTCGA 98

RESULT 13
US-09-513-999C-10082/c
; Sequence 10082, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 10082
; LENGTH: 391
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-10082

Query Match      2.2%; Score 18; DB 4; Length 391;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      129 GCGTTCCTGGCAGCAGCC 146
DB      75 GCGTTCCTGGCAGCAGCC 58

RESULT 14
US-09-252-991A-1854
; Sequence 1854, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
```

```
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1854
; LENGTH: 429
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1854

Query Match      2.2%; Score 18; DB 4; Length 429;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      739 GCCGGCTGGCGCCGCCG 756
DB      159 GCCGGCTGGCGCCGCCG 176

RESULT 15
US-09-252-991A-1154/c
; Sequence 1154, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1154
; LENGTH: 513
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1154

Query Match      2.2%; Score 18; DB 4; Length 513;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      137 GGCAGCAGGCCGCCGACC 154
DB      151 GGCAGCAGGCCGCCGACC 134
```

Search completed: August 29, 2005, 19:39:34
Job time : 245 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 29, 2005, 17:23:47 ; Search time 698 Seconds
(without alignments)
7761.717 Million cell updates/sec

Title: US-10-069-353a-7

Perfect score: 828
Sequence: 1 gctgtgccaggtggtgcacac.....cgcggaagcggctgcgtcga 828

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 7331713 seqs, 3271544945 residues

Word size : 0

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	828	100.0	80161	17	US-10-329-148A-1
2	22	2.7	849	15	US-10-156-761-925
3	22	2.7	9025608	15	US-10-156-761-1
4	20	2.4	237	11	US-09-987-899-5572
5	20	2.4	263	11	US-09-987-899-5443
6	20	2.4	290	11	US-09-987-899-5444
7	20	2.4	294	11	US-09-987-899-5442

8	20	2.4	305	11	US-09-987-899-5441	Sequence 5441, Ap
9 <td>20</td> <td>2.4<td>433<td>11<th>US-09-987-899-5571</th><th>Sequence 5571, Ap</th></td></td></td>	20	2.4 <td>433<td>11<th>US-09-987-899-5571</th><th>Sequence 5571, Ap</th></td></td>	433 <td>11<th>US-09-987-899-5571</th><th>Sequence 5571, Ap</th></td>	11 <th>US-09-987-899-5571</th> <th>Sequence 5571, Ap</th>	US-09-987-899-5571	Sequence 5571, Ap
10 <td>20</td> <td>2.4<td>601<td>20<th>US-10-425-115-101089</th><th>Sequence 101089, A</th></td></td></td>	20	2.4 <td>601<td>20<th>US-10-425-115-101089</th><th>Sequence 101089, A</th></td></td>	601 <td>20<th>US-10-425-115-101089</th><th>Sequence 101089, A</th></td>	20 <th>US-10-425-115-101089</th> <th>Sequence 101089, A</th>	US-10-425-115-101089	Sequence 101089, A
11 <td>20</td> <td>2.4<td>1164<td>17<th>US-10-369-493-33854</th><th>Sequence 33854, A</th></td></td></td>	20	2.4 <td>1164<td>17<th>US-10-369-493-33854</th><th>Sequence 33854, A</th></td></td>	1164 <td>17<th>US-10-369-493-33854</th><th>Sequence 33854, A</th></td>	17 <th>US-10-369-493-33854</th> <th>Sequence 33854, A</th>	US-10-369-493-33854	Sequence 33854, A
12 <td>20</td> <td>2.4<td>1643<td>19<th>US-10-437-963-33720</th><th>Sequence 33720, A</th></td></td></td>	20	2.4 <td>1643<td>19<th>US-10-437-963-33720</th><th>Sequence 33720, A</th></td></td>	1643 <td>19<th>US-10-437-963-33720</th><th>Sequence 33720, A</th></td>	19 <th>US-10-437-963-33720</th> <th>Sequence 33720, A</th>	US-10-437-963-33720	Sequence 33720, A
13 <td>20</td> <td>2.4<td>1652<td>18<th>US-10-425-114-35015</th><th>Sequence 35015, A</th></td></td></td>	20	2.4 <td>1652<td>18<th>US-10-425-114-35015</th><th>Sequence 35015, A</th></td></td>	1652 <td>18<th>US-10-425-114-35015</th><th>Sequence 35015, A</th></td>	18 <th>US-10-425-114-35015</th> <th>Sequence 35015, A</th>	US-10-425-114-35015	Sequence 35015, A
14 <td>20</td> <td>2.4<td>1689<td>18<th>US-10-425-114-18286</th><th>Sequence 18286, A</th></td></td></td>	20	2.4 <td>1689<td>18<th>US-10-425-114-18286</th><th>Sequence 18286, A</th></td></td>	1689 <td>18<th>US-10-425-114-18286</th><th>Sequence 18286, A</th></td>	18 <th>US-10-425-114-18286</th> <th>Sequence 18286, A</th>	US-10-425-114-18286	Sequence 18286, A
15 <td>20</td> <td>2.4<td>1896<td>20<th>US-10-425-115-140615</th><th>Sequence 140615, A</th></td></td></td>	20	2.4 <td>1896<td>20<th>US-10-425-115-140615</th><th>Sequence 140615, A</th></td></td>	1896 <td>20<th>US-10-425-115-140615</th><th>Sequence 140615, A</th></td>	20 <th>US-10-425-115-140615</th> <th>Sequence 140615, A</th>	US-10-425-115-140615	Sequence 140615, A
16 <td>20</td> <td>2.4<td>1946<td>19<th>US-10-437-963-8409</th><th>Sequence 8409, Ap</th></td></td></td>	20	2.4 <td>1946<td>19<th>US-10-437-963-8409</th><th>Sequence 8409, Ap</th></td></td>	1946 <td>19<th>US-10-437-963-8409</th><th>Sequence 8409, Ap</th></td>	19 <th>US-10-437-963-8409</th> <th>Sequence 8409, Ap</th>	US-10-437-963-8409	Sequence 8409, Ap
17 <td>20</td> <td>2.4<td>2053<td>17<th>US-10-739-930-5026</th><th>Sequence 5026, Ap</th></td></td></td>	20	2.4 <td>2053<td>17<th>US-10-739-930-5026</th><th>Sequence 5026, Ap</th></td></td>	2053 <td>17<th>US-10-739-930-5026</th><th>Sequence 5026, Ap</th></td>	17 <th>US-10-739-930-5026</th> <th>Sequence 5026, Ap</th>	US-10-739-930-5026	Sequence 5026, Ap
18 <td>20</td> <td>2.4<td>2805<td>17<th>US-10-282-122A-28732</th><th>Sequence 28732, A</th></td></td></td>	20	2.4 <td>2805<td>17<th>US-10-282-122A-28732</th><th>Sequence 28732, A</th></td></td>	2805 <td>17<th>US-10-282-122A-28732</th><th>Sequence 28732, A</th></td>	17 <th>US-10-282-122A-28732</th> <th>Sequence 28732, A</th>	US-10-282-122A-28732	Sequence 28732, A
19 <td>20</td> <td>2.4<td>2985<td>17<th>US-10-282-122A-26421</th><th>Sequence 26421, A</th></td></td></td>	20	2.4 <td>2985<td>17<th>US-10-282-122A-26421</th><th>Sequence 26421, A</th></td></td>	2985 <td>17<th>US-10-282-122A-26421</th><th>Sequence 26421, A</th></td>	17 <th>US-10-282-122A-26421</th> <th>Sequence 26421, A</th>	US-10-282-122A-26421	Sequence 26421, A
20 <td>20</td> <td>2.3<td>628<td>19<th>US-10-437-963-31900</th><th>Sequence 31900, A</th></td></td></td>	20	2.3 <td>628<td>19<th>US-10-437-963-31900</th><th>Sequence 31900, A</th></td></td>	628 <td>19<th>US-10-437-963-31900</th><th>Sequence 31900, A</th></td>	19 <th>US-10-437-963-31900</th> <th>Sequence 31900, A</th>	US-10-437-963-31900	Sequence 31900, A
21 <td>19<td>2.3<td>870<td>19<th>US-10-767-701-12680</th><th>Sequence 12680, A</th></td></td></td></td>	19 <td>2.3<td>870<td>19<th>US-10-767-701-12680</th><th>Sequence 12680, A</th></td></td></td>	2.3 <td>870<td>19<th>US-10-767-701-12680</th><th>Sequence 12680, A</th></td></td>	870 <td>19<th>US-10-767-701-12680</th><th>Sequence 12680, A</th></td>	19 <th>US-10-767-701-12680</th> <th>Sequence 12680, A</th>	US-10-767-701-12680	Sequence 12680, A
22 <td>19<td>2.3<td>902<td>20<th>US-10-425-115-47646</th><th>Sequence 47646, A</th></td></td></td></td>	19 <td>2.3<td>902<td>20<th>US-10-425-115-47646</th><th>Sequence 47646, A</th></td></td></td>	2.3 <td>902<td>20<th>US-10-425-115-47646</th><th>Sequence 47646, A</th></td></td>	902 <td>20<th>US-10-425-115-47646</th><th>Sequence 47646, A</th></td>	20 <th>US-10-425-115-47646</th> <th>Sequence 47646, A</th>	US-10-425-115-47646	Sequence 47646, A
23 <td>19<td>2.3<td>1182<td>15<th>US-10-156-761-2993</th><th>Sequence 2993, Ap</th></td></td></td></td>	19 <td>2.3<td>1182<td>15<th>US-10-156-761-2993</th><th>Sequence 2993, Ap</th></td></td></td>	2.3 <td>1182<td>15<th>US-10-156-761-2993</th><th>Sequence 2993, Ap</th></td></td>	1182 <td>15<th>US-10-156-761-2993</th><th>Sequence 2993, Ap</th></td>	15 <th>US-10-156-761-2993</th> <th>Sequence 2993, Ap</th>	US-10-156-761-2993	Sequence 2993, Ap
24 <td>19<td>2.3<td>1188<td>21<th>US-10-432-928A-8</th><th>Sequence 8, Appl1</th></td></td></td></td>	19 <td>2.3<td>1188<td>21<th>US-10-432-928A-8</th><th>Sequence 8, Appl1</th></td></td></td>	2.3 <td>1188<td>21<th>US-10-432-928A-8</th><th>Sequence 8, Appl1</th></td></td>	1188 <td>21<th>US-10-432-928A-8</th><th>Sequence 8, Appl1</th></td>	21 <th>US-10-432-928A-8</th> <th>Sequence 8, Appl1</th>	US-10-432-928A-8	Sequence 8, Appl1
25 <td>19<td>2.3<td>1418<td>21<th>US-10-432-928A-11</th><th>Sequence 11, Appl1</th></td></td></td></td>	19 <td>2.3<td>1418<td>21<th>US-10-432-928A-11</th><th>Sequence 11, Appl1</th></td></td></td>	2.3 <td>1418<td>21<th>US-10-432-928A-11</th><th>Sequence 11, Appl1</th></td></td>	1418 <td>21<th>US-10-432-928A-11</th><th>Sequence 11, Appl1</th></td>	21 <th>US-10-432-928A-11</th> <th>Sequence 11, Appl1</th>	US-10-432-928A-11	Sequence 11, Appl1
26 <td>19<td>2.3<td>1418<td>21<th>US-10-432-928A-69</th><th>Sequence 69, Appl1</th></td></td></td></td>	19 <td>2.3<td>1418<td>21<th>US-10-432-928A-69</th><th>Sequence 69, Appl1</th></td></td></td>	2.3 <td>1418<td>21<th>US-10-432-928A-69</th><th>Sequence 69, Appl1</th></td></td>	1418 <td>21<th>US-10-432-928A-69</th><th>Sequence 69, Appl1</th></td>	21 <th>US-10-432-928A-69</th> <th>Sequence 69, Appl1</th>	US-10-432-928A-69	Sequence 69, Appl1
27 <td>19<td>2.3<td>1895<td>20<th>US-10-425-115-33974</th><th>Sequence 33974, A</th></td></td></td></td>	19 <td>2.3<td>1895<td>20<th>US-10-425-115-33974</th><th>Sequence 33974, A</th></td></td></td>	2.3 <td>1895<td>20<th>US-10-425-115-33974</th><th>Sequence 33974, A</th></td></td>	1895 <td>20<th>US-10-425-115-33974</th><th>Sequence 33974, A</th></td>	20 <th>US-10-425-115-33974</th> <th>Sequence 33974, A</th>	US-10-425-115-33974	Sequence 33974, A
28 <td>19<td>2.3<td>123025<td>17<th>US-10-289-762-1</th><th>Sequence 1, Appl1</th></td></td></td></td>	19 <td>2.3<td>123025<td>17<th>US-10-289-762-1</th><th>Sequence 1, Appl1</th></td></td></td>	2.3 <td>123025<td>17<th>US-10-289-762-1</th><th>Sequence 1, Appl1</th></td></td>	123025 <td>17<th>US-10-289-762-1</th><th>Sequence 1, Appl1</th></td>	17 <th>US-10-289-762-1</th> <th>Sequence 1, Appl1</th>	US-10-289-762-1	Sequence 1, Appl1
29 <td>18<td>2.2<td>198<td>19<th>US-10-437-963-3364</th><th>Sequence 3364, Ap</th></td></td></td></td>	18 <td>2.2<td>198<td>19<th>US-10-437-963-3364</th><th>Sequence 3364, Ap</th></td></td></td>	2.2 <td>198<td>19<th>US-10-437-963-3364</th><th>Sequence 3364, Ap</th></td></td>	198 <td>19<th>US-10-437-963-3364</th><th>Sequence 3364, Ap</th></td>	19 <th>US-10-437-963-3364</th> <th>Sequence 3364, Ap</th>	US-10-437-963-3364	Sequence 3364, Ap
30 <td>18<td>2.2<td>268<td>9<th>US-09-923-876-5454</th><th>Sequence 5454, Ap</th></td></td></td></td>	18 <td>2.2<td>268<td>9<th>US-09-923-876-5454</th><th>Sequence 5454, Ap</th></td></td></td>	2.2 <td>268<td>9<th>US-09-923-876-5454</th><th>Sequence 5454, Ap</th></td></td>	268 <td>9<th>US-09-923-876-5454</th><th>Sequence 5454, Ap</th></td>	9 <th>US-09-923-876-5454</th> <th>Sequence 5454, Ap</th>	US-09-923-876-5454	Sequence 5454, Ap
31 <td>18<td>2.2<td>268<td>10<th>US-09-923-876-5454</th><th>Sequence 5454, Ap</th></td></td></td></td>	18 <td>2.2<td>268<td>10<th>US-09-923-876-5454</th><th>Sequence 5454, Ap</th></td></td></td>	2.2 <td>268<td>10<th>US-09-923-876-5454</th><th>Sequence 5454, Ap</th></td></td>	268 <td>10<th>US-09-923-876-5454</th><th>Sequence 5454, Ap</th></td>	10 <th>US-09-923-876-5454</th> <th>Sequence 5454, Ap</th>	US-09-923-876-5454	Sequence 5454, Ap
32 <td>18<td>2.2<td>273<td>9<th>US-09-923-876-5094</th><th>Sequence 5094, Ap</th></td></td></td></td>	18 <td>2.2<td>273<td>9<th>US-09-923-876-5094</th><th>Sequence 5094, Ap</th></td></td></td>	2.2 <td>273<td>9<th>US-09-923-876-5094</th><th>Sequence 5094, Ap</th></td></td>	273 <td>9<th>US-09-923-876-5094</th><th>Sequence 5094, Ap</th></td>	9 <th>US-09-923-876-5094</th> <th>Sequence 5094, Ap</th>	US-09-923-876-5094	Sequence 5094, Ap
33 <td>18<td>2.2<td>273<td>10<th>US-09-923-876-5094</th><th>Sequence 5094, Ap</th></td></td></td></td>	18 <td>2.2<td>273<td>10<th>US-09-923-876-5094</th><th>Sequence 5094, Ap</th></td></td></td>	2.2 <td>273<td>10<th>US-09-923-876-5094</th><th>Sequence 5094, Ap</th></td></td>	273 <td>10<th>US-09-923-876-5094</th><th>Sequence 5094, Ap</th></td>	10 <th>US-09-923-876-5094</th> <th>Sequence 5094, Ap</th>	US-09-923-876-5094	Sequence 5094, Ap
34 <td>18<td>2.2<td>275<td>9<th>US-09-923-876-5296</th><th>Sequence 5296, Ap</th></td></td></td></td>	18 <td>2.2<td>275<td>9<th>US-09-923-876-5296</th><th>Sequence 5296, Ap</th></td></td></td>	2.2 <td>275<td>9<th>US-09-923-876-5296</th><th>Sequence 5296, Ap</th></td></td>	275 <td>9<th>US-09-923-876-5296</th><th>Sequence 5296, Ap</th></td>	9 <th>US-09-923-876-5296</th> <th>Sequence 5296, Ap</th>	US-09-923-876-5296	Sequence 5296, Ap
35 <td>18<td>2.2<td>275<td>10<th>US-09-923-876-5296</th><th>Sequence 5296, Ap</th></td></td></td></td>	18 <td>2.2<td>275<td>10<th>US-09-923-876-5296</th><th>Sequence 5296, Ap</th></td></td></td>	2.2 <td>275<td>10<th>US-09-923-876-5296</th><th>Sequence 5296, Ap</th></td></td>	275 <td>10<th>US-09-923-876-5296</th><th>Sequence 5296, Ap</th></td>	10 <th>US-09-923-876-5296</th> <th>Sequence 5296, Ap</th>	US-09-923-876-5296	Sequence 5296, Ap
36 <td>18<td>2.2<td>276<td>9<th>US-09-923-876-5430</th><th>Sequence 5430, Ap</th></td></td></td></td>	18 <td>2.2<td>276<td>9<th>US-09-923-876-5430</th><th>Sequence 5430, Ap</th></td></td></td>	2.2 <td>276<td>9<th>US-09-923-876-5430</th><th>Sequence 5430, Ap</th></td></td>	276 <td>9<th>US-09-923-876-5430</th><th>Sequence 5430, Ap</th></td>	9 <th>US-09-923-876-5430</th> <th>Sequence 5430, Ap</th>	US-09-923-876-5430	Sequence 5430, Ap
37 <td>18<td>2.2<td>276<td>9<th>US-09-923-876-5511</th><th>Sequence 5511, Ap</th></td></td></td></td>	18 <td>2.2<td>276<td>9<th>US-09-923-876-5511</th><th>Sequence 5511, Ap</th></td></td></td>	2.2 <td>276<td>9<th>US-09-923-876-5511</th><th>Sequence 5511, Ap</th></td></td>	276 <td>9<th>US-09-923-876-5511</th><th>Sequence 5511, Ap</th></td>	9 <th>US-09-923-876-5511</th> <th>Sequence 5511, Ap</th>	US-09-923-876-5511	Sequence 5511, Ap
38 <td>18<td>2.2<td>276<td>10<th>US-09-923-876-5430</th><th>Sequence 5430, Ap</th></td></td></td></td>	18 <td>2.2<td>276<td>10<th>US-09-923-876-5430</th><th>Sequence 5430, Ap</th></td></td></td>	2.2 <td>276<td>10<th>US-09-923-876-5430</th><th>Sequence 5430, Ap</th></td></td>	276 <td>10<th>US-09-923-876-5430</th><th>Sequence 5430, Ap</th></td>	10 <th>US-09-923-876-5430</th> <th>Sequence 5430, Ap</th>	US-09-923-876-5430	Sequence 5430, Ap
39 <td>18<td>2.2<td>276<td>10<th>US-09-923-876-5511</th><th>Sequence 5511, Ap</th></td></td></td></td>	18 <td>2.2<td>276<td>10<th>US-09-923-876-5511</th><th>Sequence 5511, Ap</th></td></td></td>	2.2 <td>276<td>10<th>US-09-923-876-5511</th><th>Sequence 5511, Ap</th></td></td>	276 <td>10<th>US-09-923-876-5511</th><th>Sequence 5511, Ap</th></td>	10 <th>US-09-923-876-5511</th> <th>Sequence 5511, Ap</th>	US-09-923-876-5511	Sequence 5511, Ap
40 <td>18<td>2.2<td>285<td>9<th>US-09-923-876-5479</th><th>Sequence 5479, Ap</th></td></td></td></td>	18 <td>2.2<td>285<td>9<th>US-09-923-876-5479</th><th>Sequence 5479, Ap</th></td></td></td>	2.2 <td>285<td>9<th>US-09-923-876-5479</th><th>Sequence 5479, Ap</th></td></td>	285 <td>9<th>US-09-923-876-5479</th><th>Sequence 5479, Ap</th></td>	9 <th>US-09-923-876-5479</th> <th>Sequence 5479, Ap</th>	US-09-923-876-5479	Sequence 5479, Ap
41 <td>18<td>2.2<td>285<td>10<th>US-09-923-876-5479</th><th>Sequence 5479, Ap</th></td></td></td></td>	18 <td>2.2<td>285<td>10<th>US-09-923-876-5479</th><th>Sequence 5479, Ap</th></td></td></td>	2.2 <td>285<td>10<th>US-09-923-876-5479</th><th>Sequence 5479, Ap</th></td></td>	285 <td>10<th>US-09-923-876-5479</th><th>Sequence 5479, Ap</th></td>	10 <th>US-09-923-876-5479</th> <th>Sequence 5479, Ap</th>	US-09-923-876-5479	Sequence 5479, Ap
42 <td>18<td>2.2<td>291<td>20<th>US-10-425-115-12019</th><th>Sequence 12019, A</th></td></td></td></td>	18 <td>2.2<td>291<td>20<th>US-10-425-115-12019</th><th>Sequence 12019, A</th></td></td></td>	2.2 <td>291<td>20<th>US-10-425-115-12019</th><th>Sequence 12019, A</th></td></td>	291 <td>20<th>US-10-425-115-12019</th><th>Sequence 12019, A</th></td>	20 <th>US-10-425-115-12019</th> <th>Sequence 12019, A</th>	US-10-425-115-12019	Sequence 12019, A
43 <td>18<td>2.2<td>295<td>9<th>US-09-294-093B-2513</th><th>Sequence 2513, Ap</th></td></td></td></td>	18 <td>2.2<td>295<td>9<th>US-09-294-093B-2513</th><th>Sequence 2513, Ap</th></td></td></td>	2.2 <td>295<td>9<th>US-09-294-093B-2513</th><th>Sequence 2513, Ap</th></td></td>	295 <td>9<th>US-09-294-093B-2513</th><th>Sequence 2513, Ap</th></td>	9 <th>US-09-294-093B-2513</th> <th>Sequence 2513, Ap</th>	US-09-294-093B-2513	Sequence 2513, Ap
44 <td>18<td>2.2<td>295<td>9<th>US-09-294-093B-2996</th><th>Sequence 2996, Ap</th></td></td></td></td>	18 <td>2.2<td>295<td>9<th>US-09-294-093B-2996</th><th>Sequence 2996, Ap</th></td></td></td>	2.2 <td>295<td>9<th>US-09-294-093B-2996</th><th>Sequence 2996, Ap</th></td></td>	295 <td>9<th>US-09-294-093B-2996</th><th>Sequence 2996, Ap</th></td>	9 <th>US-09-294-093B-2996</th> <th>Sequence 2996, Ap</th>	US-09-294-093B-2996	Sequence 2996, Ap
45 <td>18<td>2.2<td>308<td>9<th>US-09-294-093B-3598</th><th>Sequence 3598, Ap</th></td></td></td></td>	18 <td>2.2<td>308<td>9<th>US-09-294-093B-3598</th><th>Sequence 3598, Ap</th></td></td></td>	2.2 <td>308<td>9<th>US-09-294-093B-3598</th><th>Sequence 3598, Ap</th></td></td>	308 <td>9<th>US-09-294-093B-3598</th><th>Sequence 3598, Ap</th></td>	9 <th>US-09-294-093B-3598</th> <th>Sequence 3598, Ap</th>	US-09-294-093B-3598	Sequence 3598, Ap

ALIGNMENTS

RESULT 1
US-10-329-148A-1
; Sequence 1, Appl1
; Publication No. US20040023343A1
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OR INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50469 DIV1
; CURRENT APPLICATION NUMBER: US/10/329,148A
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/603,207B
; PRIOR FILING DATE: EARLIER FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/370,700
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 80161
; TYPE: DNA
; ORGANISM: Saccharopolyspora spinosa
US-10-329-148A-1
Query Match 100.0%; Score 828; DB 17; Length 80161;
Best Local Similarity 100.0%; Pred. No. 0;

```
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTGGCCAGTGGCGCACCAATCGACGAGTTGGGACAGATGATGACCTGGGACG 60
DB 20168 GTGTGGCCAGTGGCGCACCAATCGACGAGTTGGGACAGATGATGACCTGGGACG 20227
QY 61 CCGTTGCTGAATCGGTGCGGGGCGGCGCCCTTGCCCATCCACACGAGCTACTGGAGAAC 120
DB 20228 CCGTTGCTGAATCGGTGCGGGGCGGCGCCCTTGCCCATCCACACGAGCTACTGGAGAAC 20287
QY 121 GACGGGGCGGGGCTTCTTGAGAGAGCGCGGACCGGCTTCAACGACCTTGTGCGCGAAGG 180
DB 20288 GACGGGGCGGGGCTTCTTGAGAGAGCGCGGACCGGCTTCAACGACCTTGTGCGCGAAGG 20347
QY 181 ACCGTGCTGATGAGGCGGCTTGAGACTGCTGATGTGGGGTGGGTAACCGGACCAAGCG 240
DB 20348 ACCGTGCTGATGAGGCGGCTTGAGACTGCTGATGTGGGGTGGGTAACCGGACCAAGCG 20407
QY 241 CTGGCGGTGCGCGCGACCAACCGGATCCAGATCCCGGATCAACGCTCAGCCAGGTGCA 300
DB 20408 CTGGCGGTGCGCGCGACCAACCGGATCCAGATCCCGGATCAACGCTCAGCCAGGTGCA 20467
QY 301 GTGGCCATGCGCGGCTGATGGCGGACCGGAACTPAAGGCAACGGGTGAGACTTTCTG 360
DB 20468 GTGGCCATGCGCGGCTGATGGCGGACCGGAACTPAAGGCAACGGGTGAGACTTTCTG 20527
QY 361 TGCGTGAATGCCATGTCCCTGCGGTACCGGACCAATGCTTTGACGCGCGCTGGGCCATG 420
DB 20528 TGCGTGAATGCCATGTCCCTGCGGTACCGGACCAATGCTTTGACGCGCGCTGGGCCATG 20587
QY 421 CAGTGGCTGTGAGATGTCCGAACCGGACCGGTCCATCCGGGAAATCTTGAAGTACTC 480
DB 20588 CAGTGGCTGTGAGATGTCCGAACCGGACCGGTCCATCCGGGAAATCTTGAAGTACTC 20647
QY 481 AAACCCGGGTGATCCTCGGCGTCAACGAGTGTCAAAACGAGGCGGCGCGGCGGATG 540
DB 20648 AAACCCGGGTGATCCTCGGCGTCAACGAGTGTCAAAACGAGGCGGCGCGGCGGATG 20707
QY 541 CCGGTGCTCGGGGAGAGGTGCGGACCGGCTTCCGATCTGCTGGCTGAGCACTTCTG 600
DB 20708 CCGGTGCTCGGGGAGAGGTGCGGACCGGCTTCCGATCTGCTGGCTGAGCACTTCTG 20767
QY 601 GAATGCTGCTGCAAGCGGGGTTTCAGATCTCGATTGGGAGAGCTGTCTGAGAGACC 660
DB 20768 GAATGCTGCTGCAAGCGGGGTTTCAGATCTCGATTGGGAGAGCTGTCTGAGAGACC 20827
QY 661 CGGTACTTCATGCGCGAAGTTGCGGAGAGCTGCTGGGCAACAGCAGGGAATGCGGAC 720
DB 20828 CGGTACTTCATGCGCGAAGTTGCGGAGAGCTGCTGGGCAACAGCAGGGAATGCGGAC 20887
QY 721 AGGTACGGGCGCGGCTGTGCGCGGCTGGGCGCGGCTCTGCGATTATGAGAAATATGCC 780
DB 20888 AGGTACGGGCGCGGCTGTGCGCGGCTGGGCGCGGCTCTGCGATTATGAGAAATATGCC 20947
QY 781 CACGACATGGGCTTATGCGATTTCGACGGGCGGGAAGCGGCTGCGCTGA 828
DB 20948 CACGACATGGGCTTATGCGATTTCGACGGGCGGGAAGCGGCTGCGCTGA 20995

RESULT 2
US-10-156-761-925
; Sequence 925, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
```

```
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 925
; LENGTH: 849
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) .. (849)
US-10-156-761-925

Query Match 2.7%; Score 22; DB 15; Length 849;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 AGCCCGCCGACCGGCTCACCGA 164
DB 149 AGCCCGCCGACCGGCTCACCGA 170

RESULT 3
US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match 2.7%; Score 22; DB 15; Length 9025608;
Best Local Similarity 100.0%; Pred. No. 0.097;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 AGCCCGCCGACCGGCTCACCGA 164
DB 1137332 AGCCCGCCGACCGGCTCACCGA 1137311

RESULT 4
US-09-987-899-5572
; Sequence 5572, Application US/09987899
; Publication No. US2004011682A1
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Nordin
; APPLICANT: Liu, Jindong
; APPLICANT: Miller, Phillip W.
```

```
; APPLICANT: O Connell, Keith M.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated
; FILE REFERENCE: 16517.258
; CURRENT APPLICATION NUMBER: US 09/987,899
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: US 09/262,979
; PRIOR FILING DATE: 1999-03-04
; PRIOR APPLICATION NUMBER: US 60/076,712
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 7341
; SEQ ID NO 5572
; LENGTH: 237
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB143-052-Q1-E1-E4
US-09-987-899-5572

Query Match          2.4%; Score 20; DB 11; Length 237;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 492 CATCTCGGCGTCACCGAG 511
DB 192 CATCTCGGCGTCACCGAG 211

RESULT 5
US-09-987-899-5443
; Sequence 5443, Application US/09987899
; Publication No. US20040116682A1
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Nordine
; APPLICANT: Liu, Jindong
; APPLICANT: Miller, Phillip W.
; APPLICANT: O Connell, Keith M.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated
; FILE REFERENCE: 16517.258
; CURRENT APPLICATION NUMBER: US/09/987,899
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: US 09/262,979
; PRIOR FILING DATE: 1999-03-04
; PRIOR APPLICATION NUMBER: US 60/076,712
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 7341
; SEQ ID NO 5443
; LENGTH: 263
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700041823H1
US-09-987-899-5443

Query Match          2.4%; Score 20; DB 11; Length 263;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 492 CATCTCGGCGTCACCGAG 511
DB 162 CATCTCGGCGTCACCGAG 181

RESULT 6
US-09-987-899-5444
; Sequence 5444, Application US/09987899
; Publication No. US20040116682A1
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Nordine
; APPLICANT: Liu, Jindong
; APPLICANT: Miller, Phillip W.
; APPLICANT: O Connell, Keith M.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated
```

```
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated
; FILE REFERENCE: 16517.258
; CURRENT APPLICATION NUMBER: US/09/987,899
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: US 09/262,979
; PRIOR FILING DATE: 1999-03-04
; PRIOR APPLICATION NUMBER: US 60/076,712
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 7341
; SEQ ID NO 5444
; LENGTH: 290
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700094321H1
US-09-987-899-5444

Query Match          2.4%; Score 20; DB 11; Length 290;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 492 CATCTCGGCGTCACCGAG 511
DB 240 CATCTCGGCGTCACCGAG 259

RESULT 7
US-09-987-899-5442
; Sequence 5442, Application US/09987899
; Publication No. US20040116682A1
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Nordine
; APPLICANT: Liu, Jindong
; APPLICANT: Miller, Phillip W.
; APPLICANT: O Connell, Keith M.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated
; FILE REFERENCE: 16517.258
; CURRENT APPLICATION NUMBER: US/09/987,899
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: US 09/262,979
; PRIOR FILING DATE: 1999-03-04
; PRIOR APPLICATION NUMBER: US 60/076,712
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 7341
; SEQ ID NO 5442
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700214146H1
US-09-987-899-5442

Query Match          2.4%; Score 20; DB 11; Length 294;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 492 CATCTCGGCGTCACCGAG 511
DB 249 CATCTCGGCGTCACCGAG 268

RESULT 8
US-09-987-899-5441
; Sequence 5441, Application US/09987899
; Publication No. US20040116682A1
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Nordine
; APPLICANT: Liu, Jindong
; APPLICANT: Miller, Phillip W.
; APPLICANT: O Connell, Keith M.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated
```



```
/ TITLE OF INVENTION: With the Carbon Assimilation Pathway
/ FILE REFERENCE: 16517.258
/ CURRENT APPLICATION NUMBER: US/09/987,899
/ CURRENT FILING DATE: 2001-11-16
/ PRIOR APPLICATION NUMBER: US 09/262,979
/ PRIOR FILING DATE: 1999-03-04
/ PRIOR APPLICATION NUMBER: US 60/076,712
/ PRIOR FILING DATE: 1998-03-06
/ NUMBER OF SEQ ID NOS: 7341
/ SEQ ID NO 5441
/ LENGTH: 305
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: 700095671H1
US-09-987-899-5441

Query Match          2.4%; Score 20; DB 11; Length 305;
Best Local Similarity 100.0%; Pred.No. 3.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      492 CATCTCGGCGTCACCGAG 511
DB      202 CATCTCGGCGTCACCGAG 221
|||||
|||||

RESULT 9
US-09-987-899-5571
/ Sequence 5571, Application US/09987899
/ Publication No. US20040116682A1
/ GENERAL INFORMATION:
/ APPLICANT: Cheikh, Nordine
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Miller, Phillip W.
/ APPLICANT: O Connell, Keith M.
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated
/ TITLE OF INVENTION: With the Carbon Assimilation Pathway
/ FILE REFERENCE: 16517.258
/ CURRENT APPLICATION NUMBER: US/09/987,899
/ CURRENT FILING DATE: 2001-11-16
/ PRIOR APPLICATION NUMBER: US 09/262,979
/ PRIOR FILING DATE: 1999-03-04
/ PRIOR APPLICATION NUMBER: US 60/076,712
/ PRIOR FILING DATE: 1998-03-06
/ NUMBER OF SEQ ID NOS: 7341
/ SEQ ID NO 5571
/ LENGTH: 433
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3079-007-Q1-X1-B3
US-09-987-899-5571

Query Match          2.4%; Score 20; DB 11; Length 433;
Best Local Similarity 100.0%; Pred.No. 3.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      492 CATCTCGGCGTCACCGAG 511
DB      237 CATCTCGGCGTCACCGAG 256
|||||
|||||

RESULT 10
US-10-425-115-101089
/ Sequence 101089, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants
```

```
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 101089
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MNT4577_23706C.1
US-10-425-115-101089

Query Match          2.4%; Score 20; DB 20; Length 601;
Best Local Similarity 100.0%; Pred.No. 3.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      363 CGTCATGCCATGTCCTGC 382
DB      481 CGTCATGCCATGTCCTGC 500
|||||
|||||

RESULT 11
US-10-369-493-33854
/ Sequence 33854, Application US/10369493
/ Publication No. US2003023675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xiandeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 33854
/ LENGTH: 1164
/ TYPE: DNA
/ ORGANISM: magnetite-containing magnetic coccus
US-10-369-493-33854

Query Match          2.4%; Score 20; DB 17; Length 1164;
Best Local Similarity 100.0%; Pred.No. 3.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      92 GCGCATCCACACGCGCTAC 111
DB      14 GCGCATCCACACGCGCTAC 33
|||||
|||||

RESULT 12
US-10-437-963-33720
/ Sequence 33720, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
```

```
/ SEQ ID NO 33720
/ LENGTH: 1643
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_37804C.1
US-10-437-963-33720

Query Match
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2.4%; Score 20; DB 19; Length 1643;

QY 492 CATCTCGGCGTCACCGAG 511
DB 135 CATCTCGGCGTCACCGAG 154

RESULT 13
US-10-425-114-35015
/ Sequence 35015, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 35015
/ LENGTH: 1652
/ TYPE: DNA
/ ORGANISM: Zea mays subsp. mexicana
/ FEATURE:
/ OTHER INFORMATION: Clone ID: UC-ZMFLTEOSINTE008F02_FLI
US-10-425-114-35015

Query Match
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2.4%; Score 20; DB 18; Length 1652;

QY 492 CATCTCGGCGTCACCGAG 511
DB 233 CATCTCGGCGTCACCGAG 252

RESULT 14
US-10-425-114-18286
/ Sequence 18286, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 18286
/ LENGTH: 1689
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
```

```
/ OTHER INFORMATION: Clone ID: LIB3079-016-D7_FLI
US-10-425-114-18286

Query Match
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2.4%; Score 20; DB 18; Length 1689;

QY 492 CATCTCGGCGTCACCGAG 511
DB 260 CATCTCGGCGTCACCGAG 279

RESULT 15
US-10-425-115-140615
/ Sequence 140615, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5322)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 140615
/ LENGTH: 1896
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_59724C.1
US-10-425-115-140615

Query Match
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2.4%; Score 20; DB 20; Length 1896;

QY 492 CATCTCGGCGTCACCGAG 511
DB 361 CATCTCGGCGTCACCGAG 380

Search completed: August 29, 2005, 19:51:16
Job time : 716 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 29, 2005, 18:38:52 ; Search time 4680 Seconds
(without alignments)
2847.262 Million cell updates/sec

Title: US-10-069-353a-8

Perfect score: 1444
Sequence: 1 VLPGAPTSQGVQGMVLT.....YKXNDMGAILITARKPVG 275

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-O/cgpn2.1/USPTO.spool/US10069353/runac.26082005.172053.14666/app.query.fasta_1.455
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LDOFCL=0 -LDOPEXT=0
-UNITS=bits8 -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcio -NORM=ext -HEAPSIZB=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10069353 @CGN 1.1 3731 @runac.26082005.172053.14666 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MMAP TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

1: gb ha:*
2: gb htg:*
3: gb ln:*
4: gb om:*
5: gb ov:*
6: gb pac:*
7: gb ph:*
8: gb pl:*
9: gb pr:*
10: gb ro:*
11: gb stc:*
12: gb sy:*
13: gb un:*
14: gb vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No. Query Match Length DB ID Description

1	1444	100.0	828	6	AX089422	Sequence
2	1444	100.0	45624	6	AX089419	Sequence
3	1444	100.0	50000	6	AX089416	Sequence
4	1444	100.0	80161	1	AY007564	Saccharop

SUMMARIES

5	1444	100.0	80161	6	AR165018	Sequence
6	1444	100.0	80161	6	BD137649	BioSynthe
7	1444	100.0	80161	6	AR281866	Sequence
8	1315	91.1	36538	6	AX600587	Sequence
9	441.5	30.6	5998	1	AB032524	Streptomy
10	441.5	30.6	302675	1	AP005024	Streptomy
11	439.5	30.4	36538	6	AX600587	Sequence
12	439	30.4	1410	6	E07846	DNA sequence
13	438.5	30.4	846	6	AX598613	Sequence
14	438.5	30.4	52101	6	AX598593	Sequence
15	435.5	30.2	900	6	AX089434	Sequence
16	435.5	30.2	45624	6	AX089419	Sequence
17	435.5	30.2	50000	6	AX089416	Sequence
18	435.5	30.2	80161	1	AY007564	Saccharop
19	435.5	30.2	80161	6	AR165018	Sequence
20	435.5	30.2	80161	6	BD137649	BioSynthe
21	435.5	30.2	80161	6	AR281866	Sequence
22	401.5	27.8	109528	1	AF040570	Amycolato
23	379	26.2	6085	6	AF03004	Sequence
24	379	26.2	6085	6	AR144763	Sequence
25	376.5	26.1	852	6	BD269407	Mitomycin
26	376.5	26.1	852	6	AR266893	Sequence
27	376.5	26.1	53500	6	BD269447	Mitomycin
28	376.5	26.1	53500	6	AR266933	Sequence
29	376.5	26.1	63734	1	AF127374	Streptomy
30	375.5	26.0	17512	1	AB071405	Lechevall
31	375.5	26.0	25681	1	SAE41459	Saccharot
32	375.5	26.0	26144	1	AB090952	Lechevall
33	375.5	26.0	28654	1	AF534707	Streptomy
34	373.5	25.9	22999	1	AB088119	Streptomy
35	372.5	25.8	846	1	AY396042	Streptomy
36	361.5	25.0	30000	6	AX250261	Sequence
37	361.5	25.0	103450	1	AF440781	Streptomy
38	345	23.9	828	6	BD269406	Mitomycin
39	343.5	23.8	12381	6	AR266892	Sequence
40	343.5	23.8	12381	6	BD217469	Polyketid
41	343.5	23.8	12381	6	AX006889	Sequence
42	340	23.5	12637	1	AF323753	Streptomy
43	329.5	22.8	21064	1	AB088224	Streptomy
44	320.5	22.2	2122	6	AF03001	Sequence
45	320.5	22.2	2122	6	AR144762	Sequence

ALIGNMENTS

RESULT 1
AX089422 828 bp DNA linear PAT 21-MAR-2001
LOCUS Sequence 7 from Patent WO0116303.
DEFINITION AX089422
ACCESSION AX089422.1 GI:13443683
VERSION
KEYWORDS
SOURCE
ORGANISM
Saccharopolyspora spinosa
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Pseudonocardineae; Pseudonocardaceae; Saccharopolyspora.
REFERENCE
1 Ederz, G., Moehle, V., Froede, R., Velten, R. and Salas, J.A.
AUTHORS Nucleic acids which code for the enzyme activities of the spinosyn
biosynthesis
TITLE Patent: WO 0116303-A 7 08-MAR-2001;
JOURNAL BAYER AG (DE)
FEATURES
source
1. 828
Location/Qualifiers
/organism="Saccharopolyspora spinosa"
/mol_type="unassigned DNA"
/db_xref="taxon:60894"
1. 828
/note="unnamed protein product; ORF1: O-Methyltransferase"
/codon_start=1
/transl_table=11
/protein_id="CAC34874.1"
/db_xref="GI:13443684"

translation="MLPGAPTSQGVQWYDVLTPILNSVAGSPCALHHGWENDGARA
SMOQADRLTDLVAERTVLDGVRLLDVCCTGGOPALRLARPAENALQIGITVSQVOVA
IADQARERLSHRVDFSCVDAMSLPYPNADAMAPMAEEMSPDAIIEIRLV
KRGILGTEVVRKAGGMPVSGDRMPTGLRICIABOLIESLRAGEILLDMEEVVS
RTRVFMPPQAEELAAHQHIDIRYGPNAVAGMAAACVDIEKVAHMDNGVAILTRKEVGS"

ORIGIN

Alignment Scores:
Pred. No.: 1,86e-100 Length: 828
Score: 1444.00 Matches: 275
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-069-353a-8 (1-275) x AX089422 (1-828)

QY 1 ValLeuProGlyValAlaProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThr 20
DB 1 GTGTTGCCAGGTGGCGCACCAATCGCAGGTGGGCAAGATGATGACCTGGTCAAG 60
QY 21 ProLeuLeuAsnSerValAlaGlyGlyProGlyProGlyAlaLeuHisGlyTyrTrpGluAsn 40
DB 61 CCGTTGCTAACTCGGTCCGGCGGAGCCCTTGCCATCACCAGGCTACTGGAGAAC 120
QY 41 AspGlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
DB 121 GACGGGCGGGCTTTCCTGGCAGCAGGCGCCGACCGGCTCACGACCTTGCCCGAACGG 180
QY 61 ThrValLeuAspGlyValAlaArgLeuLeuAspValGlyGlyGlyThrGlyGlnProAla 80
DB 181 ACCGTGCTCGATGGCGGCTTCGACTGCTCGATGGGGGTGGCGTAACCGCACCAACGAGG 240
QY 81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGln 100
DB 241 CTGCGGCTGCGCGGACCAACCGCATCCAGATCACCGGATCACCGTCAAGCGAGTGCAC 300
QY 101 ValAlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120
DB 301 GTGGCCATGCGCGCTGATGGCGCACCGGACCGGACTAAGCCACCGGGTGGACTTCTCG 360
QY 121 CysValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
DB 361 TCCGTGCAATGCATGCTCCGTGCGTAACCGGACCAATGCTTTCGACGCCCTCGGGCCATG 420
QY 141 GlnSerLeuLeuGluMetSerGluProAspArgAlaIleArgGluIleLeuArgValLeu 160
DB 421 CAGTGCCTTGGAGATGTCGAACCGGACCGTCCATCCGGGAATCTTTCGAGTATCTC 480
QY 161 LysProGlyGlyIleLeuGlyValAlaThrGluValAlaValArgGluAlaGlyGlyMet 180
DB 481 AAACCCGGGTGCATCTCGGGCTCACCGAGTCTGTCAAACGAGAACGGCGCGCGGAGTGG 540
QY 181 ProValSerGlyAspArgTrpProThrArgIleLeuArgIleCysLeuAlaGluGlnLeu 200
DB 541 CCGGCTGTCGGGACAGGTGGCGGACCGGCTTCGATCTGCTCGTGAAGCAACTTCTCG 600
QY 201 GlnSerLeuArgAlaAlaGlyPheGluIleLeuAspTrpGluAspValSerSerArgThr 220
DB 601 GAATGCTGCTGCGACGGGGTTTCGAGATCTTCGATTGGAGAGACGTGCTGCGAGAAC 660
QY 221 ArgTyrPheMetProGlnPheAlaGluGluLeuAlaAlaHisGlnHisGlyIleAlaAsp 240
DB 661 CCGTACTTCATGCGCAGATTGCCCGAAGAGCTCGCTGGGCACACAGACGGAGTCCGGAGC 720
QY 241 ArgTyrGlyProAlaValAlaGlyTyrTrpAlaAlaAlaValCysAspTyrGlnLysTyrAla 260
DB 721 AGTACCGGGCCCGCTGCTCCGCTGGCGGCGCGGCTCTGCGATTATGAGAAATATATGCC 780
QY 261 HisAspMetGlyTyrAlaIleLeuThrAlaArgLysProValGly 275
DB 781 CACGCAATGGGCTATGCGATCTCGACGGCGCGAAGCCGGTGGC 825

RESULT 2
AX089419/c 45624 bp DNA linear PAT 21-MAR-2001
Sequence 4 from Patent WO0116303.
ACCESSION AX089419
VERSION AX089419.1 GI:13443680
KEYWORDS
SOURCE
ORGANISM
Saccharopolyspora spinosa
Saccharopolyspora spinosa
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Pseudonocardineae; Pseudonocardiceae; Saccharopolyspora.

REFERENCE
AUTHORS
Eberz, G., Moehrle, V., Froede, R., Velten, R. and Salas, J.A.
Nucleic acids which code for the enzyme activities of the spinosyn
biosynthesis

JOURNAL Patent: WO 0116303-A 4 08-MAR-2001;

FEATURES
source location/Qualifiers
1. 45624
/organism="Saccharopolyspora spinosa"
/mol_type="unassigned DNA"
/db_xref="taxon:60894"

ORIGIN

Alignment Scores:
Pred. No.: 1.57e-98 Length: 45624
Score: 1444.00 Matches: 275
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-069-353a-8 (1-275) x AX089419 (1-45624)

QY 1 ValLeuProGlyValAlaProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThr 20
DB 828 GTGTTGCCAGGTGGCGCACCAATCGCAGGTGGGCAAGATGATGACCTGGTCAAG 769
QY 21 ProLeuLeuAsnSerValAlaGlyGlyProGlyProGlyAlaLeuHisGlyTyrTrpGluAsn 40
DB 768 CCGTTGCTAACTCGGTCCGGCGGAGCCCTTGCCATCACCAGGCTACTGGAGAAC 709
QY 41 AspGlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
DB 708 GACGGGCGGGCTTCTCGGACAGGCGCCGACCGGCTCACCGGCTTGTGCGGAAACG 649
QY 61 ThrValLeuAspGlyValAlaArgLeuLeuAspValGlyGlyGlyThrGlyGlnProAla 80
DB 648 ACCGTGCTCGATGGCGGCTTCGACTGCTCGATGGGGGTGGCGTAACCGGACCAACGAGG 589
QY 81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGln 100
DB 588 CTGCGGCTGCGCGGACCAACCGCATCCAGATCCACCGGCACTGACCGCAAGTGCAC 529
QY 101 ValAlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120
DB 528 GTGGCCATGCGCGCTGATGGCGCACCGGACCGGACTAAGCCACCGGGTGGACTTCTCG 469
QY 121 CysValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
DB 468 TCCGTGCAATGCATGCTCCCTCGGACCGGACCAATGCTTTCGAGCGCGCTGGGCCATG 409
QY 141 GlnSerLeuLeuGluMetSerGluProAspArgAlaIleArgGluIleLeuArgValLeu 160
DB 408 CAGTGCCTTGGAGATGTCGAACCGGACCGTCCATCCGGGAATCTTTCGAGTATCTC 349
QY 161 LysProGlyGlyIleLeuGlyValAlaThrGluValAlaValArgGluAlaGlyGlyMet 180
DB 348 AAACCCGGGTGCATCTCGGGCTCACCGAGGTCGTCAAACGAGAACGGCGCGCGGAGTGG 289
QY 181 ProValSerGlyAspArgTrpProThrArgIleLeuArgIleCysLeuAlaGluGlnLeu 200
DB 288 CCGGTGTCGGGACAGGTGGCGCACCGGCTTCGAGATCTGCTGCGTGAACCACTTCTCG 229

QY 201 GUSerLeuArgAlaIaIaIaGlyPheGluIleLeuAspTrpGluAspValSerSerArgThr 220
DB 228 GAATCGCTCGTCACACGGGGTTTCGAGATCTTCGATTGGAGGACGGTCTGAGAGACC 169
QY 221 ArgTyPheMetProGlnPheAlaGluGluLeuAlaAlaHisGlnIleGlyIleAlaAsp 240
DB 168 CGGTACTTCATGCGCGAGTTCCGCGAAGAGCTCTGCTGCCACGACGCGGATCCGGGAC 109
QY 241 ArgTyGlyProAlaValAlaIaIaGlyTrpAlaAlaIaValCysAspTrpGluValGly 260
DB 108 AGGTACGGGCGGCTGTCGCGCGCTGGGCGCGCGCTGCGATTATGAAATATGCG 49
QY 261 HisAspMetGlyTyTrpAlaIleLeuThrAlaArgLysProValGly 275
DB 48 CACGACATGGGCTATGCGATTCTGACGGCGCGGAGCGCGGTCCGCG 4
RESULT 3
AX089416 AX089416 50000 bp DNA linear PAT 21-MAR-2001
LOCUS Sequence 1 from Patent WO0116303.
ACCESSION AX089416
VERSION AX089416.1 GI:13443677
KEYWORDS
SOURCE Saccharopolyspora spinosa
ORGANISM Saccharopolyspora spinosa
Bacteria; Actinobacteridae; Actinomycetales;
Pseudonocardineae; Pseudonocardaceae; Saccharopolyspora.
REFERENCE
AUTHORS Bherg, G., Moehle, V., Froede, R., Veltien, R. and Salas, J. A.
TITLE Nucleic acids which code for the enzyme activities of the spinosyn
JOURNAL biosynthesis
Patent: WO 0116303-A 1 08-MAR-2001;
BAYER AG (DE)
FEATURES
source location/Qualifiers
1..50000
/organism="Saccharopolyspora spinosa"
/mol_type="unassigned DNA"
/db_xref="taxon:60894"
ORIGIN
Alignment Scores:
Pred. No.: 1.74e-98 Length: 50000
Score: 1444.00 Matches: 275
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-10-069-353A-8 (1-275) x AX089416 (1-50000)
QY 1 ValLeuProGlyGlyAlaProThrSerGlnGlnValGlyGlnMetTyAspLeuValThr 20
DB 44797 GTGTGGCAGGTGGCGACCAACATCCACAGAGTTGGGAGAGTATGACTGTGTACG 44856
QY 21 ProLeuLeuAsnSerValAlaIaIaGlyProCysAlaIleHisGlyTyTrpGluAsn 40
DB 44857 CCGTGTGTAACCTGCTCGCGCGCGCGCCCTGCGCCATCCACACGCGCTACTGCGAGAAC 44916
QY 41 AspGlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
DB 44917 GACCGCGCGGCTCTCTGCGACGACGCGCGACCGCGCTCACCGACTTGTGCGGACGG 44976
QY 61 ThrValLeuAspGlyGlyValArgLeuLeuAspValGlyCysGlyThrGlyGlnProAla 80
DB 44977 ACCGTGCTCATGGCGGCGGTTCGACTGCTGATGTGGGGTGGCGTACCGGACACGACGG 45036
QY 81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGln 100
DB 45037 CTGCGCGTCTGCGCGCAACGCGCATCCAGATCCCGGCAACACGCTCAGCGAGTGCAC 45096
QY 101 ValAlaIleAlaIaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120

DB 45097 GTGGCCATGCGCGCTGATTGCGCACGCGAACCAGCACTAAGCCACCGGCTGACTTCG 45156
QY 121 CysValAspAlaMetSerLeuProTyTrpAspAsnAlaPheAspAlaAlaTrpAlaMet 140
DB 45157 TCGTGTGATGCAATGCTCTGCGCGTACCGGACCAATGCTTTCGACGCCCTCGGGCCATG 45216
QY 141 GlnSerLeuGlnGlnMetSerGlnProAspArgAlaIleArgGluIleLeuArgValLeu 160
DB 45217 CAGTCCCTTGTGGAATGTCGACACCGGACCGGTGCATCCGGGAATCTTCGAGTATTC 45276
QY 161 LysProGlyGlyIleLeuGlyValThrGluValIleValArgGluAlaGlyGlyMet 180
DB 45277 AAACCGGTGCATCTCTCGCGCTCACCGAGTGTGCAACGAGACCGGCGCGGAGTG 45336
QY 181 ProValSerGlyAspArgTrpProThrArgLysLeuArgIleCysLeuAlaGluGlnLeu 200
DB 45337 CCGGTGTCCGGGACAGGTGGCGACCGGCTTCGATGCTGCTGACTAGCACTTCTG 45396
QY 201 GUSerLeuArgAlaIaIaIaGlyPheGluIleLeuAspTrpGluAspValSerSerArgThr 220
DB 45397 GAATCGCTCGTCACACGGGGTTCCAGATCTTCGATTGGGAGAGCGTGTCTGAGGAC 45456
QY 221 ArgTyPheMetProGlnPheAlaGluGluLeuAlaAlaHisGlnIleGlyIleAlaAsp 240
DB 45457 CGGTACTTCATGCGCGAGTTCCGCGAAGAGCTGCTGCCACACGACGCGGATTCGCGGAC 45516
QY 241 ArgTyGlyProAlaValAlaIaIaGlyTrpAlaAlaIaValCysAspTrpGluValGly 260
DB 45517 AGGTACGGGCGCGCTGTCGCGCGCTGGGCGCGCGCTGCGATTATGAAATATGCG 45576
QY 261 HisAspMetGlyTyTrpAlaIleLeuThrAlaArgLysProValGly 275
DB 45577 CACGACATGGGCTATGCGATTCTGACGGCGCGGAGCGCGGTCCGCG 45621
RESULT 4
AY007564 AY007564 80161 bp DNA linear BCT 16-MAY-2001
LOCUS Saccharopolyspora spinosa probable lyxR-type transcriptional
DEFINITION regulator, probable keto acyl reductase, spinosid biosynthetic gene
cluster, complete sequence, and probable exodeoxyribonuclease V
genes, complete cds; and unknown gene.
ACCESSION AY007564.1 GI:13162633
VERSION
KEYWORDS
SOURCE Saccharopolyspora spinosa
ORGANISM Saccharopolyspora spinosa
Bacteria; Actinobacteridae; Actinomycetales;
Pseudonocardineae; Pseudonocardaceae; Saccharopolyspora.
REFERENCE
AUTHORS 1 (bases 1 to 80161)
Waldron, C., Matsushima, P., Rosteck, P. R., Jr., Broughton, M. C.,
Turner, D., Madduri, K., Crawford, K. P., Merlo, D. J. and Baltz, R. H.
TITLE Cloning and analysis of the spinosid biosynthetic gene cluster of
Saccharopolyspora spinosa
JOURNAL Chem. Biol. 8 (5), 487-499 (2001)
MEDLINE 21257765
PUBMED 11358695
REFERENCE 2 (bases 1 to 80161)
Waldron, C., Matsushima, P., Rosteck, P. R., Jr., Broughton, M. C.,
Turner, D., Madduri, K., Crawford, K. P., Merlo, D. J. and Baltz, R. H.
TITLE Direct Submision
JOURNAL Submitted (18-AUG-2000) Dow Agrosciences, 9330 Zionsville Rd.,
Indianapolis, IN 46268, USA
FEATURES
source location/Qualifiers
1..80161
/organism="Saccharopolyspora spinosa"
/mol_type="genomic DNA"
/db_xref="taxon:60894"
complement(1135..1971)
/note="ORF L16"
/codon_start=1
/transl_table=11
/product="probable lyxR-type transcriptional regulator"
/protein_id="AAG23282.1"

gene /db xref="GI:13162654"
/translation="METRELRYFVAVAEELHFGRAAORLGLIAQPLSRITIAQLERGL
VYLQRTSRKVSULTTEGAMLLTEGALIGALNRTORPATISQPSVLAAKASG
ELAKLIDVAEPGAVVDLIGSOPKTHDORADVALIHOPDPPTAEIDITLN
TEQOYAILPTSHPLASERHVRADVSSIPDLPLAKRPGDGYPRGSEVRNQQLP
QMLAGRTTVHPSSRVNLLEGLAIVFLDAPDVTIVIAMPPHSRSLALGLVAVAT
IL"
2024. .2791
/note="ORF 115"
/codon_start=1
/transl_table=11
/product="probable keto acyl reductase"
/protein_id="AAG23281.1"
/db xref="GI:13162653"
/translation="WLVRRPRPATLESMEQOTIALVTGANKIGIETIAAGLALGWS
VIGARDHORGEDVAVKLRADGVDAFAVSLDVTDOASVAAAAALEERAGRLDVVNN
AGIAGWPEEPTVTPASIRAVETNVIGVAVTNAMPLIPRSRPRITVNSSVAS
LTLQTTPEVDLGGISGAYSPSKTFILNALTIOYAKELSDTNIKINNACGCVATDLNGF
HGTSTPADGARIAIRLATLPDDGPTGGMFDGANVPW"
complement(3416. .4165)
/gene="spnS"
complement(3416. .4165)
/gene="spnS"
/note="probably involved in forosamine biosynthesis"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="probable N-dimethyltransferase"
/protein_id="AAG23280.1"
/db xref="GI:13162652"
/translation="MSRVSDTFPAETSSVYSPDHADIYALHSARGDMAAGEVYQL
VTRLPBAQSLDVAACGTGAHLERPRAYAKVAGLESDAMKEIIRIVPEVPIHIGD
IRDFLGEFPDVTICLCTFAAYMTVDLRLVTRMMARHLAPGVAVIEPWFPKF
IDGVGAVAHHEGVERISRLSHSVEGRSRMVRVTVAAPEICIRPTEFELISLPTED
EXTALDAGIRAEYLPGAPNGRGLFVGRN"
complement(4168. .5325)
/gene="spnR"
complement(4168. .5325)
/gene="spnR"
/note="probably involved in forosamine biosynthesis"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="probable aminotransferase"
/protein_id="AAG23279.1"
/db xref="GI:13162651"
/translation="MINLHOPITGTEELDAIEVPSNWIGGPRTPPEAFPAHILG
VDPEVFLNCGTAAFLITVOYLDLGRPDVTLPSISVAAANAIASSGARPFCDVD
PRTLPRLDVARAITPATKAVLLHYGSGPGEVTAIDPCREKGLMLIEDSACVAVS
SVHGTACGTFGDLATWSPDAMKILVTGCGMFTADPELHARARLAYHGLQMSGFD
SAKSNRMWDIREDIGORLIGNDTAALGSQLKLEPINRREIATQYDRLISDV
GVLLPPTLPDGHVSHFYVYQALAPRDVOAQQMLERGIYTSRYRPPLEKPIYRA
DCKLPASADACRCTLPLHPBLDABEVRYADERQAKVHHISGRSLRK"
complement(5363. .6751)
/gene="spnQ"
complement(5363. .6751)
/gene="spnQ"
/note="probably involved in forosamine biosynthesis"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="probable NDP-hexose-3,4-dehydratase"
/protein_id="AAG23278.1"
/db xref="GI:13162650"
/translation="MOSRKTALGKGARVMSCDTTCATTEMVPDAKDRILIASVDY
HREQSPTPVAGSTPIRPSGAVLDEDDVVALVEALERIRIAGNARPFSEFARFEG
LRKAHLVNSGSSANLALUSLTSPLUGARLKPGEVITTAIVGFTTINPVOVNLVP
FVVDVLGYNAIPRIRIKAAVETRRATMLATTIGNPRADEIETIAKEHELFLVEDN
CDVAGSTRGRLTGFGDLITVSPFAHHTISGEGCVLITGLIELARIIEELRIGRD
CWCEPVNTCKRKFEDHGLTLPFGYDHKYTSSHVGNMKTTDLQALATLSOLSKISA
FGSARRMMRRLREGLSGIPGLLIPVATPHSPSPRGALITISADAGCTRAALVNFLE
SRNICTRLIFGNGITRHPAFBOYRIRIADALTNDSITVDRTFVGVYFGITDQMDIYV
VESTIAEFVAKSS"

gene 7083. .8450
/gene="spnR"
CDS 7083. .8450
/gene="spnR"
/note="involved in forosamine addition"
/codon_start=1
/transl_table=11
/product="probable NDP-forosamyltransferase"
/protein_id="AAG23277.1"
/db xref="GI:13162649"
/translation="MLIGMLPGCSIALGEMRVLTPLPASSHRFNVLPLAMLRAG
HEVVAICENNVSVTGAGLTAVPVGEBLDTLSIAAKNELVLSGVSPDEKGRBELP
DELSINSGRDTPDAVEQLHLVDRLSDLDFGFAKQWPDVLVWAMVCSGVVARALG
ARHARMVALDMLKSLRSGFLEYOESKPEQRVDPDGLTWGAKAKLAKGAFDEBIVTG
QATIDPIPSMWRLPVLDLYISMRFVYNGPVALPEMLERPTKRVCTITRGLTRLS
RVTEYOESQVSDHMAVERLIRGVARLDVEITALSDDEVEMGELPSNVHVEVPL
NELLESGSVIIHNSSTTOETATNAGPQOLIPGTMPDESRARLADRGAGLVDPA
TFTEDDVRGQARLRLDEPSPANALIRREIESPSPHDIYVRLKLVARENRRTQ
SDGRP"
8967. .10427
/gene="spnO"
CDS 8967. .10427
/gene="spnO"
/note="involved in forosamine biosynthesis"
/codon_start=1
/transl_table=11
/product="probable NDP-hexose-2,3-dehydratase"
/protein_id="AAG23276.1"
/db xref="GI:13162648"
/translation="WSSSVEAEASAAAPLGSNNTRRPVDSALSNCMITPTTEFCHL
ADRGENSFETNRJLPDRLSKWKFDASTENLVHADGPFVTEGLQVETNGAASWQ
PIINQAEVGLIGLIVKEIDVGLHLMKAPKPGVNVNQLQSPVQATRSNVTQAHGS
VPVPLDVLGSGRPLVVDVLOSQSGMFPKRRKRNNAVVEQEEVPVLPFCMLTGG
VVALRDQNTVMOTRTVLSGIPPHDSATGELLAASHEPQAVARLSHGDSSSIS
BAVGMPEAKARVRLRATRVPLSVDMVWSTTDEIAHODKPAVIVITVASATREVA
SWTPMILPBEQGEITALIVKRGIVGLVGHVARIEAKYKTAEIAPVQSVANQGT
PSNDMPFLDVLTADEPTVYESILSEBGRFYQAOENRRIIEHEDFAARPPSDR
WMITGQUGELIRSTHFLNIQARSLVASHSIMALGR"
10436. .11434
/gene="spnM"
CDS 10436. .11434
/gene="spnM"
/note="involved in forosamine biosynthesis"
/codon_start=1
/transl_table=11
/product="probable NDP-hexose-3-ketoreductase"
/protein_id="AAG23275.1"
/db xref="GI:13162647"
/translation="WRRPVRIIGVLGCASFAMRMLPAMCDAETEVAVASRDBAKAE
RFAARFCEAVLGYQRLLERPDIDAVVPPPLPGMHAEMIGKALEADKHVLAKEPLTTT
ASDPAARLVGLARKNLLRENVYELHHGRHDVVNDLQSGEIGELREFTVFGIPPLP
PTDILRYTELEGGGLLIDIGVAPAAAGHFLGLPTVGVASHVESQESVILGSLG
SEGCTVHILGCGFHHYRAVELMGSPRTLVVPPAPPEMAQVITIEKGVYDELS
LPABDQVRKAVTAPARDIRAGTCVDDPAVAGDSSEMTIQALALVEALIGQARRCGST"
complement(11530. .12492)
/gene="spnM"
complement(11530. .12492)
/gene="spnM"
/note="involved in spinosyn aglycone biosynthesis"
/codon_start=1
/transl_table=11
/product="SpnM"
/protein_id="AAG23274.1"
/db xref="GI:13162646"
/translation="MPAAVSGTVLVPIPMWRERDRLITTEFAVGNHGLSGOVAPEYLAR
TGTEPBEELIANVALDRGMAYVITDYBGLGPRGTHITTVYGAQAGNAMDARAARLGG
SGLTITDPCVGIWGAQGGQASAPFGEHLPTIYAPLRRAAAGAVPTDLDITIRKNDG
VFTCPVAVGLAVGHAAGVLDLPEFDELTEAGRTADVOREIGAPELVIRFGRELSDFL
DTSLEFQOPTRARLALASVAGRNGPVPVTLVYHSTIDEIVPFAFGERLDSYRAAGT
PVRRHPISGIAHPAALASRRVAVSWFDEHSEBSAISGRDAN"
complement(12696. .13547)
/gene="spnL"
CDS complement(12696. .13547)

```
/gene="spnL"
/notes="involved in spinosyn aglycone biosynthesis;
possibly involved in polyketide bridging"
/codon_start=1
/transl_table=11
/product="methyltransferase-like protein"
/protein_id="A02327.1"
/db_xref="GI:13162645"
/translations="MESTFDALHGRPLHNGWAGREDACATPMSDAADLTDLFI
DKAALPRAHLFDLGGNGQPVVRAACAGSVRTGTITVAQHILAAATRLANETGLGS
LEFDLVDGQQLPYDGFPOAAMQSVQVVDQAAIREVHRIIEBGRFVLDITR
VRLPEEYAAVMTGTATHTLSPALVSEGEFLEVTDLTAOTRCWMSYVDELLKRL
DELAGEPAAVGTGQRYLGDIAKKHPPQAQILAAVAEYRKHPDVARMESEFWLL
QARKKS"
complement(13592..14785)
/gene="spnK"
complement(13592..14785)

Alignment Scores:
Pred. No.: 2,93e-98 Length: 80161
Score: 1444.00 Matches: 275
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-10-069-353a-8 (1-275) x AY007564 (1-80161)

QY 1 ValLeuProGlyValAlaProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThr 20
Db GTGTTGCGAGTGGCGGCGACCAACATCGCAGAGTTGGCGAGTGTATGACCTGGTACG 20227
QY 21 ProLeuLeuAsnSerValAlaGlyProCyAspAlaIleHisGlyTyrTrpGluAsn 40
Db CCGTTGCTGAACCTCGTCCGCGGCGCCCTCGGCCATCCACACCGGCTACTGGGAGAAC 20287
QY 41 AspGlyValArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
Db GACGGGCGGGCTTCCTGCGACGAGCGCGGACCGGCTGACCCGACTTGTGCGCGAACGG 20347
QY 61 ThrValLeuAspGlyValAlaArgLeuLeuAspValGlyCyAspGlyThrGlyGlnProAla 80
Db ACCGTGCTGCATGGCGGCGCTTGCATCTGCATGTGGGGTGGCGTACCGGACCAACGCG 20407
QY 81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyTyrThrValSerGlnValGln 100
Db CTGGCGCTCGCGCGGCGACCAACGCGATCCAGATCCCGGCAATCACCGTCAAGGAGTCAA 20467
QY 101 ValAlaIleAlaAlaAspCyAspAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120
Db GTGGCCATCGCCGCTGATTCGCGACGCGAACGCGGACTAAGCCACCGGGTGGACTTCTCG 20527
QY 121 CysValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
Db TGGCTGCATGCCATGTCCCTGCCCTACCCGGAACAATCTTTGACGCGCGCTGGGCATG 20587
QY 141 GlnSerLeuLeuGlnMetSerGlnProAspArgAlaIleArgGlnIleLeuArgValLeu 160
Db CAGTCGCTGTTGAGATGTCGGAACCGGACCGTCCATCCGGGAAATCTTTCAGATCTC 20647
QY 161 LysProGlyValIleLeuGlyValThrGluValValLysArgGluAlaGlyValGlyMet 180
Db AAACCCGATGGCATCCCTCGGCGTACCGAGGTCTCTCAACCGAAGAGCGGGCGGGGATG 20707
QY 181 ProValSerGlyAspArgTrpProThrGlyLeuArgIleCysLeuAlaGlnGlnLeu 200
Db CCGGTCGCGGGGACAGGTGCCACCGGCTTGGATTCGCTCGGCTGAGCAATCTCTG 20767
QY 201 GluSerLeuArgAlaAlaGlyPheGlnIleLeuAspTrpGluAspValSerSerArgThr 220
Db GAATCGCTGCGTGCAGCGGGGTTGAGATCTTCGATTGGGAGGACGCGTGTCTGAGGACC 20827
QY 221 ArgTyrThrMetProGlnPheAlaGluGluLeuAlaAlaHisGlnHisGlyIleAlaAsp 240
```

```
Db 20828 CCGTACTTCATGCCCGACGATTCGCCGAAGAGCTCGCTGGCGACCAACGCGGATCCCGGAC 20887
QY 241 ArgTyrGlyProAlaValAlaGlyTrpAlaAlaAlaValCysAspTyrGluArgTyrAla 260
Db 20888 AGGTACGGGCGCGGCTGTCCCGCGCGGCGCGCGCTGGCATTAAGAAATATGCGC 20947
QY 261 HisAspMetGlyTyrAlaIleLeuThrAlaArgLysProValGly 275
Db 20948 CACGACATGGGCTATGCGATTCGACGCGCGCGGAACCGGTGCGC 20992

RESULT 5
LOCUS AR165018 80161 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6274350.
ACCESSION AR165018
VERSION AR165018.1 GI:16238344
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 80161)
AUTHORS Balcz,R.H., Crawford,K.P., Broughton,M.Christine., Madduri,K.,
Merlo,D.J., Turner,J.R., Treadway,P.J. and Waldron,C.
TITLE Biosynthetic genes for spinosyn insecticide production
JOURNAL Patent: US 6274350-A 1 14-AUG-2001;
FEATURES
Location/Qualifiers
1..80161
/organism="unknown"
/mol_type="unasigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 2,93e-98 Length: 80161
Score: 1444.00 Matches: 275
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-069-353a-8 (1-275) x AR165018 (1-80161)

QY 1 ValLeuProGlyValAlaProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThr 20
Db GTGTTGCGAGTGGCGGCGACCAACATCGCAGAGTTGGCGAGTGTATGACCTGGTACG 20227
QY 21 ProLeuLeuAsnSerValAlaGlyProCyAspAlaIleHisGlyTyrTrpGluAsn 40
Db CCGTTGCTGAACCTCGTCCGCGGCGCCCTCGGCCATCCACACCGGCTACTGGGAGAAC 20287
QY 41 AspGlyValArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
Db GACGGGCGGGCTTCCTGCGACGAGCGCGGACCGGCTGACCCGACTTGTGCGCGAACGG 20347
QY 61 ThrValLeuAspGlyValAlaArgLeuLeuAspValGlyCyAspGlyThrGlyGlnProAla 80
Db ACCGTGCTGCATGGCGGCGCTTGCATCTGCATGTGGGGTGGCGTACCGGACCAACGCG 20407
QY 81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyTyrThrValSerGlnValGln 100
Db CTGGCGCTCGCGCGGCGACCAACGCGATCCAGATCCCGGCAATCACCGTCAAGGAGTCAA 20467
QY 101 ValAlaIleAlaAlaAspCyAspAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120
Db AAACCCGATGGCATCCCTCGGCGTACCGAGGTCTCTCAACCGAAGAGCGGGCGGGGATG 20527
QY 121 CysValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
Db TGGCTGCATGCCATGTCCCTGCCCTACCCGGAACAATCTTTGACGCGCGCTGGGCATG 20587
QY 141 GlnSerLeuLeuGlnMetSerGlnProAspArgAlaIleArgGlnIleLeuArgValLeu 160
Db CAGTCGCTGTTGAGATGTCGGAACCGGACCGTCCATCCGGGAAATCTTTCAGATCTC 20647
QY 161 LysProGlyValIleLeuGlyValThrGluValValLysArgGluAlaGlyValGlyMet 180
Db AAACCCGATGGCATCCCTCGGCGTACCGAGGTCTCTCAACCGAAGAGCGGGCGGGGATG 20707
QY 181 ProValSerGlyAspArgTrpProThrGlyLeuArgIleCysLeuAlaGlnGlnLeu 200
Db CCGGTCGCGGGGACAGGTGCCACCGGCTTGGATTCGCTCGGCTGAGCAATCTCTG 20767
QY 201 GluSerLeuArgAlaAlaGlyPheGlnIleLeuAspTrpGluAspValSerSerArgThr 220
Db GAATCGCTGCGTGCAGCGGGGTTGAGATCTTCGATTGGGAGGACGCGTGTCTGAGGACC 20827
QY 221 ArgTyrThrMetProGlnPheAlaGluGluLeuAlaAlaHisGlnHisGlyIleAlaAsp 240
```


QY 161 LysProGlyGlyIleuenglyValThrGluValIleuArgGluAlaGlyGlyMet 180
Db 20648 AAACCCGGGCGATCTCCGGCTCACCGAGTCGTCAACGAGAACGGGGCGGATG 20707
QY 181 ProValSerGlyAspArgTrpProThrglyLeuArgIleCysLeuAlaGluGlnLeu 200
Db 20708 CCGGTGTCCGGGAGACAGGTGGCCGACCCGGCTTCGGATCTGCTGGCTGACACTTCTG 20767
QY 201 GluSerLeuArgAlaAlaGlyPheGluIleuAspTrpGluAspValSerSerArgThr 220
Db 20768 GAATGCTGCTGACGCGGGGTTCAGATCTCGATTGGAGAGACGTGTGTGAGAGACC 20827
QY 221 ArgTyrPheMetProGlnPheAlaGluGlnLeuAlaAlaHisGlnHisGlyIleAlaAsp 240
Db 20828 CGGTACTTCATGCGCGAGTTCCCGAAGAGCTCGCTGGCCACACAGACGGGATCCGGAC 20887
QY 241 ArgTyrGlyProAlaValAlaGlyTyrTrpAlaAlaAlaValCysAspTyrGluLysTyrAla 260
Db 20888 AGGTACGGGCGCGCTGTCCCGCGGTGGCGCGCGCGCTCTCGATTATGAGAAATATGCG 20947
QY 261 HisAspMetGlyTyrAlaIleLeuThrAlaArgLysProValGly 275
Db 20948 CACGACATGGCTATGCGATTCTGACGCGCGAGAACCGGTGCGC 20992

RESULT 6
LOCUS BD137649 80161 bp DNA linear PAT 18-SEP-2002
DEFINITION Biosynthetic gene for spinosyn insecticide production.
ACCESSION BD137649
VERSION BD137649.1 GI:23232594
KEYWORDS JP 2002505881-A/1.
SOURCE Saccharopolyspora spinosa
ORGANISM Saccharopolyspora spinosa
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Pseudonocardiales; Pseudonocardaceae; Saccharopolyspora.
REFERENCE 1 (bases 1 to 80161)
AUTHORS Baltz,R.H., Broughton,M.C., Crawford,K.P., Madduri,K., Merlo,D.J.,
Treadway,P.J., Turner,J.R. and Waldron,C.
TITLE Biosynthetic genes for spinosyn insecticide production
JOURNAL Patent: JP 2002505881-A 1 26-FEB-2002;
DOW AGRSCIENCES LLC
COMMENT OS Saccharopolyspora spinosa
PN JP 2002505881-A/1
PD 26-FEB-2002
PI 16-FEB-1999 JP 2000535754
PR 09-MAR-1998 US 09/036987
PI RICHARD H BALTZ,M CHRISTINE BROUGHTON,KATHRYN P CRAWFORD,PI
KRISHNAMURTHY MADDURI,DONALD J MERLO,PATTI J TREADWAY,JAN R PI
TURNER,
PI CLIVE WALDRON
PC C12N15/09,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12P19/62,C12Q1/
PC 68//
PC C07K14/195,C12N15/00,C12N5/00
CC Biosynthetic genes for spinosyn insecticide production FH
Key Location/Qualifiers
FT source 1. 80161
FT /organism='Saccharopolyspora spinosa'.
FEATURES
source Location/Qualifiers
1. 80161
/organism='Saccharopolyspora spinosa'
/mol_type='genomic DNA'
/db_xref='taxon:60894'

ORIGIN
Alignment Scores:
Pred. No.: 2.93e-98 Length: 80161
Score: 1444.00 Matches: 275
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
Db:

US-10-069-353A-8 (1-275) x BD137649 (1-80161)
QY 1 ValLeuProGlyGlyAlaProThrSerGlnGlnValGlyGlyMetTyrAspLeuValThr 20
Db 20168 GTGTGGCCAGGTGGGACCAACATCGCAGAGGTGGGCGAGATGATGACTGTGACG 20227
QY 21 ProLeuLeuAsnSerValAlaGlyGlyProCysAlaIleHisGlyTyrTrpGluAsn 40
Db 20228 CGGTGCTGAACTGGGTGGGGGGGGCCCTGGCCATCCACCAACGGCTTACGGGAGAAC 20287
QY 41 AspGlyValAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
Db 20288 GACGGGCGGGCTCTCTGGCACAGCGCCGCCACCGGCTCACCGACCTTGTGCCGAACGG 20347
QY 61 ThrValLeuAspGlyGlyValArgLeuLeuAspValGlyCysGlyThrGlyGlnProAla 80
Db 20348 ACCGTGCTCATGGCGGGCTTCGACTGCTCATGTGGGGTCCGGTACCGAACACAGCG 20407
QY 81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGln 100
Db 20408 CTGGCGGTGGCGCGGACCAACCGCATCCAGATCACCGGCATCACCGTACGCGCAGGTGCA 20467
QY 101 ValAlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120
Db 20468 GTGGCCATCGCGCTGATTGGGCAACGCAACGCGAGCTAACCCACCGGTTGACTTCTCG 20527
QY 121 CysValaAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
Db 20528 TGGTCGATGCCATGTCCCTCCCTGACCCGACAAATGCTTTGACGCCGCTGGAGCAATG 20587
QY 141 GluSerLeuLeuGluMetSerGluProAspArgAlaIleArgGluIleLeuArgValLeu 160
Db 20588 CAGTCGCTGTTGGAGATGTCGAAACCGGACCGTGCATCCGGAAATCTTCGATCTC 20647
QY 161 LysProGlyGlyIleuenglyValIleThrglyValValIleArgGluAlaGlyGlyMet 180
Db 20648 AAACCCGGGTGCATCTCTCGGGTCAACGAGGTCTCTCAACGAGAACGGGGCGGATG 20707
QY 181 ProValSerGlyAspArgTrpProThrglyLeuArgIleCysLeuAlaGluGlnLeu 200
Db 20708 CCGGTGTCCGGGAGACAGGTGGCCGACCCGGCTTCGGATCTGCTGGCTGAGCAACTTCTG 20767
QY 201 GluSerLeuArgAlaAlaGlyPheGluIleLeuAspTrpGluAspValSerSerArgThr 220
Db 20768 GAATGCTGCTGACGCGGGGTTCAGATCTCTCGATTGGAGAGACGTGTGTGAGAGACC 20827
QY 221 ArgTyrPheMetProGlnPheAlaGluGlnLeuAlaAlaHisGlnHisGlyIleAlaAsp 240
Db 20828 CCGTACTTCATGCGCGAGTTCCCGAAGAGCTCGCTGGCCACACAGCGGATGCGGAC 20887
QY 241 ArgTyrGlyProAlaValAlaGlyTyrTrpAlaAlaAlaValCysAspTyrGlyLysTyrAla 260
Db 20888 AGGTACGGGCGCGCTGTCCCGCGGTGGCGCGCGCTCTCGATTATGAGAAATATGCG 20947
QY 261 HisAspMetGlyTyrAlaIleLeuThrAlaArgLysProValGly 275
Db 20948 CACGACATGGCTATGCGATTCTGACGCGCGGAGAACCGGTGCGC 20992

RESULT 7
LOCUS AR281866 80161 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1 from patent US 6521406.
ACCESSION AR281866
VERSION AR281866.1 GI:29717767
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 80161)
AUTHORS Baltz,R.H., Broughton,M.C., Crawford,K.P., Madduri,K., Merlo,D.J.,
Treadway,P.J., Turner,J.R. and Waldron,C.
TITLE SpnG, a gene for spinosyn insecticide biosynthesis
JOURNAL Patent: US 6521406-A 1 18-FEB-2003;

FEATURES
 source
 Location/Qualifiers
 1..80161
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Alignment Scores:
 Pred. No.: 2,93e-98 Length: 80161
 Score: 1444.00 Matches: 275
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-10-069-353a-8 (1-275) x AR281866 (1-80161)

QY 1 ValLeuProGlyVgIyAlaProThrsSerGlnGlnValGlyGlnMetYrAspLeuValThr 20
 |||||
 Db 20168 GTGTTCCAGGTGGCGGCAACCAATCGACAGCTTGGCGAGTATGACCTGCTCAGC 20227
 |||||

QY 21 ProLeuLeuAsnSerValAlaGlyProCyAlaIleHisHisGlyYrTTPGluAan 40
 |||||
 Db 20228 CCGTTCTGAACCTCGGTCGGCGGCGCCCTGCGCCATCCACAGGCTACTGGAGAAC 20287
 |||||

QY 41 AspGlyArgAlaSerTTPGlnGlnAlaAspArgLeuThrAspLeuValAlaGluArg 60
 |||||
 Db 20288 GACGGCGGGCTTCTTGCGACAGCGCGCGACCGGCTCACCACTTGTGCGCGAACGG 20347
 |||||

QY 61 ThrValLeuAspGlyVgIyValArgLeuLeuAspValGlyCyArgYrThrGlyGlnProAla 80
 |||||
 Db 20348 ACCGTGCTGATGCGCGGCTTCATGCTGATGTGGGTGGGTACCGGACCAACGCG 20407
 |||||

QY 81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyYrThrValSerGlnValGln 100
 |||||
 Db 20408 CTGCGGCTCGCGCGGCAACCGGATCCAGATCACCGGCAATCCGTCACAGGTGCA 20467
 |||||

QY 101 ValAlaIleAlaAlaAspCyAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120
 |||||
 Db 20468 GTGGCCATCCCGCTGATTCGCGACCGGAACTAGCCACCGGCTGACTTCTCG 20527
 |||||

QY 121 CyValAlaAspAlaMetSerLeuProYrProAspAsnAlaPheAspAlaATrPalamet 140
 |||||
 Db 20528 TGGTCATGCTCATGTCCTGCGGACCGGACATCTTTCACCGCGCTGGCCATG 20587
 |||||

QY 141 GlnSerLeuLeuGlnMetSerGluProAspArgAlaIleArgGluIleLeuArgValLeu 160
 |||||
 Db 20588 CAGTCGCTGTGGAGATGTCGAACCGGACCGGCAATCTTCCAGTACTC 20647
 |||||

QY 161 LysProGlyGlyIleLeuGlyValThrGluValValLysArgGluArgGlyGlyMet 180
 |||||
 Db 20648 AAACCCGGTGCATCTCGCGCTGACCGAGGTCTCAACGAGAGCGGCGGGGATG 20707
 |||||

QY 181 ProValSerGlyAspArgTTPProThrGlyLeuArgIleCysLeuAlaGlnLeuLeu 200
 |||||
 Db 20708 CCGGTGTCGGGGGACAGGTGCGACCGGCTTTCGATCTGCTGAGCAACTTCTG 20767
 |||||

QY 201 GluSerLeuArgAlaAlaGlyPheGluIleLeuAspTTPGluAspValSerSerArgThr 220
 |||||
 Db 20768 GAATCGCTGGGTGAGGGGGTTCGAGATCTTCATTTGGAGGAGCGTGTGAGGACC 20827
 |||||

QY 221 ArgTyrPheMetProGlnPheAlaGluGluLeuAlaAlaHisGlnHisGlyIleAlaAsp 240
 |||||
 Db 20828 CCGTACTTCAATGCGCGAGTTCGCGAAGAGCTCTCTCGCACCGGAGGAGTTCGCGAC 20887
 |||||

QY 241 ArgTyrGlyProAlaValAlaGlyTTPAlaAlaAlaValCysAspYrGluLysYrAla 260
 |||||
 Db 20888 AGGTACGGGCGGCTGTGCGCGGCTGGCGCGCGCGCTGCTGCAATTGAGAAATATGCC 20947
 |||||

QY 261 HisAspMetGlyYrAlaIleLeuThrAlaArgLysProValGly 275
 |||||
 Db 20948 CACACATGGGCTATGCGATTCTACGGCGGGAAGCGGTCGCGC 20992
 |||||

RESULT 8

AX600587/C
 LOCUS AX600587 36538 bp DNA linear PAT 14-FEB-2003
 DEFINITION Sequence 2 from Patent WO02079477.
 ACCESSION AX600587
 VERSION AX600587.1 GI:28400303
 KEYWORDS
 SOURCE
 ORGANISM
 Saccharopolyspora sp. NRRL 30141
 Saccharopolyspora sp. NRRL 30141
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Pseudonocardineae; Pseudonocardiales; Saccharopolyspora.

REFERENCE
 1 Hahn, D.R., Jackson, J.D., Bullard, B.S., Gustafson, G.D., Waldron, C. and Mitchell, J.C.
 TITLE Biosynthetic genes for butenyl-spirosyn insecticide production
 JOURNAL Patent: WO 02079477-A 2 10-OCT-2002;
 Dow Agrosciences LLC (US)

FEATURES
 source
 Location/Qualifiers
 1..36538
 /organism="Saccharopolyspora sp. NRRL 30141"
 /mol_type="unassigned DNA"
 /db_xref="taxon:221471"
 /note="LM107129"

ORIGIN

Alignment Scores:
 Pred. No.: 7.04e-89 Length: 36538
 Score: 1315.00 Matches: 252
 Percent Similarity: 94.18% Conservative: 7
 Best Local Similarity: 91.64% Mismatches: 16
 Query Match: 91.07% Indels: 0
 DB: Gaps: 0

US-10-069-353a-8 (1-275) x AX600587 (1-36538)

QY 1 ValLeuProGlyVgIyAlaProThrsSerGlnGlnValGlyGlnMetYrAspLeuValThr 20
 |||||
 Db 938 GTGTTCCAGGTGGCGGCAACCAATCGACAGCTTGGCGAGTATGACCTGCTCAGC 879
 |||||

QY 21 ProLeuLeuAsnSerValAlaGlyProCyAlaIleHisHisGlyYrTTPGluAan 40
 |||||
 Db 878 CCGTTCTGAACCTCGGTCGGCGGCGCCCTGCGCCATCCACAGGCTACTGGAGAAC 819
 |||||

QY 41 AspGlyArgAlaSerTTPGlnGlnAlaAspArgLeuThrAspLeuValAlaGluArg 60
 |||||
 Db 818 GACGGCGGGCTTCTTGCGACAGCGCGCGACCGGCTCACCACTTGTGCGCGAACGG 759
 |||||

QY 61 ThrValLeuAspGlyVgIyValArgLeuLeuAspValGlyCyArgYrThrGlyGlnProAla 80
 |||||
 Db 758 ACCGCTGCTGATGCGCGGCTTCATGCTGATGTGGGTGGGTACCGGACCAACGCG 699
 |||||

QY 81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyYrThrValSerGlnValGln 100
 |||||
 Db 698 CTGCGGCTCGCGGCGCAACCGGATCCGAGTCAACCGGATCCGCTACGCGAGTGCAC 639
 |||||

QY 101 ValAlaIleAlaAlaAspCyAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120
 |||||
 Db 638 GCTGCATGCTCTTATGTCGCGCGGACCGGACGAGTCAACGACCAAGTGTGACTTCTG 579
 |||||

QY 121 CyValAlaAspAlaMetSerLeuProYrProAspAsnAlaPheAspAlaATrPalamet 140
 |||||
 Db 578 TGGTCGAGCGCATGTCCTGCGGCTGACCGGACCAATGCGTTCGACCTGCTGGCCATG 519
 |||||

QY 141 GlnSerLeuLeuGlnMetSerGluProAspArgAlaIleArgGluIleLeuArgValLeu 160
 |||||
 Db 518 CAGTCTGCTGAGATGTCGGAACCGGACCGGATCCGGAATAATCGTTCAAGTGTCTC 459
 |||||

QY 161 LysProGlyGlyIleLeuGlyValThrGluValValLysArgGluArgGlyGlyMet 180
 |||||
 Db 458 AAACCCGGTGCATCTCGAGTTCACCGAGTCTCAACGAGAAATAGGACGGGAGTTC 399
 |||||

QY 181 ProValSerGlyAspArgTTPProThrGlyLeuArgIleCysLeuAlaGlnLeuLeu 200
 |||||
 Db 398 CCGGTGTCGGGGGACAGGTGCGACCGGCTTTCGATCTGCTGAGCAACTTCTG 339
 |||||

Db 5381 GAATGCGCTTTCCGACGGGCTCTTCAAGCGGCGCTGGAGTCTCTCGAC 5322

Qy 146 MetSerGluProAspArgAlaIleArgIleLeuArgValLeuArgProGlyIle 165

Db 5321 ATGCCAGCCCGCACAGGTGATCCGGAGATCGCCGGGTGCTCGCCCGCGCGCG 5262

Qy 166 LeuGlyValThrGluValValLeuArg-GluAlaGlyGlyGlyMetProValSerGlyAs 185

Db 5261 CTGGCCGCTGACGACGATCGCATCTGCGCGCTTGGCGGACCGGACATGAAGCGCGGAG 5202

Qy 185 pArgTrpProThrGlyLeuArgIle-----GlyLeuAlaGluGlnLeuGlyIle 202

Db 5201 TGCACCTCCACAGTCTCGCGGCTCCGCGCCCTGCTGACATGACAGATGACCGGAAAG 5142

Qy 202 rLeuArgAlaAlaGlyPheGlyIleLeuArgProGlyAspValSerArgThrArgTy 222

Db 5141 -ATCGCGGACGCGCGGCTGATCGATGATGACATGACATGCGGATCAGGTC----- 5088

Qy 222 rPheMetProGlnPheAla-----GluGlyLeuAlaAlaGlnIleGlyIleAl 239

Db 5087 -GTGCGCCCTCTTTCGCGCGCGCTGCTGACACGTGAACGACACCTCGACGATGACG 5029

Qy 239 AspArgGlyTrgIyrProAlaValAlaGlyTrp-----AlaAlaAlaValCysAspTyrgI 257

Db 5028 GCGCGCCCTTGGAGATGCGCGCTGCGGAGATGCGGAAAGTGTGTCACAGTGCACGAGCT 4969

Qy 257 uLygTyraIleAspMetGlyTyraIleLeuThrAlaArgPro 273

Db 4968 CCCCTGACGCGCGGACATGCGCTATGCGTGCAGCGCGCGCGCGCG 4920

RESULT 10
AP005024/c 302675 bp DNA linear BCT 10-MAY-2003
LOCUS Streptomyces avermitilis genomic DNA, complete genome, section 4/30
DEFINITION

ACCESSION AP005024 BA000030
VERSION AP005024.1 GI:29604389
KEYWORDS
SOURCE Streptomyces avermitilis MA-4680
ORGANISM Streptomyces avermitilis MA-4680
Bacteria; Actinobacteria; Actinomycetaceae; Streptomyces.

REFERENCE 1 Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., Takahashi, C., Shinose, M., Takahashi, Y., Horikawa, H., Nakazawa, H., Osone, T., Kikuchi, H., Shiba, T., Sakaki, Y. and Hattori, M.
Genome sequence of an industrial microorganism Streptomyces avermitilis: deducing the ability of producing secondary metabolites
Proc. Natl. Acad. Sci. U.S.A. 98 (21), 12215-12220 (2001)

JOURNAL MEDLINE 21477403
PUBMED 11572948

TITLE 2 Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M. and Omura, S.
Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis
Nat. Biotechnol. 21 (5), 526-531 (2003)

JOURNAL MEDLINE 22608306
PUBMED 12692562

REFERENCE 3 (bases 1 to 302675)
Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., Takahashi, C., Shinose, M., Takahashi, Y., Horikawa, H., Nakazawa, H., Osone, T., Kishida, N., Director-General of Biotechnology Center, Shiba, T., Sakaki, Y. and Hattori, M.
Direct Submission
Submitted (29-MAR-2002) Director-General of Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center, 2-49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
(E-mail: bioente.go.jp, URL: http://www.bio.ente.go.jp/, Tel: 81-3-3481-1933, Fax: 81-3-3481-8424)

COMMENT This work was done in collaboration with Haruo Ikeda(*1), Jun Ishikawa(*2), Akihito Hanamoto(*3), Chigusa Takahashi(*3), Mayumi

FEATURES

source

Shinose(*3), Hiroshi Horikawa(*4), Hidekazu Nakazawa(*4), Tomomi Osone(*4), Norihiro Kishida(*4), Hisashi Kikuchi(*4), Tadayoshi Shiba(*5), Yoshiyuki Sakaki(*6,*7), Masahira Hattori(*1,*7) and Satoishi Omura(*1,*3).
Final finishing process and all annotation were done by H. Ikeda and J. Ishikawa.
*1 Kitasato Institute for Life Sciences, Kitasato University
*2 National Institute of Infectious Diseases
*3 The Kitasato Institute
*4 National Institute of Technology and Evaluation
*5 School of Science, Kitasato University
*6 Institute of Medical Science, University of Tokyo
*7 RIKEN, Genomic Sciences Center
Following url is also available.
http://avermitilis.is.kitasato-u.ac.jp.
Location/Qualifiers

1..302675
/organism="Streptomyces avermitilis MA-4680"
/mol_type="genomic DNA"
/strain="MA-4680"
/db_xref="taxon:227882"
/note="This strain is also named as strain: ATCC 31267, NCIMB 12804 or NRRL B165.-synonym: Streptomyces avermitilis"
12..425
/note="SAV750"
/codon_start=1
/transl_table=11
/product="putative TetR-family transcriptional regulator"
/protein_id="BAC68460.1"
/db_xref="GI:29604390"
/translation="MABLCIDREPARTAAKQALSGIAGGQTVARNIDAVALLPFGSG PMSGLATMRPAAARIRERIALVDGAPGTAIOEATGYIKARGLGRVAETDPRIVA LAIVTAHLLMTSPGADPFRPHARVALVD5"
1504..2592
/note="SAV751"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC68461.1"
/db_xref="GI:29604391"
/translation="MRTGVEBELLLVDPETGEPKALSTAVLARAQVDPDDEFEKE LRGOMLEFATPOTMAALGAEIVRCRKAABAGACAVALLATSPVPSILMN RYQWMAEYGTAMOEGLTGGCHVAVASDEGVAVDVIRIPMLPVALLANSPEM QKDSSTYESTYRKRWRKRPASPTFELPSPRYHYNOVAMVATGYLDGMYFLDAR SHRYTVEIRVADVCLHPDTAQLIATLARGLVESARDRARGRPEDSHVSLRLAAR QARGLSGELAHPTMRLRAESVVRALLHEVGDALDADGLDLAHEAVAEILERN GARVRELELRGSLRDYTECVROTQG"
complement(3011..3400)
/note="SAV752"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC68462.1"
/db_xref="GI:29604392"
/translation="MQTASTVSLTADVAAQAFILAHGTYVQAADGPAFLTRAD ADVILARGTOLPPEQRDRRAAGIILATLADGLHDEKRLREGVITMPLEREP WGERLEQITDPNGVILIQFVEMATSDPT"
complement(3457..3648)
/note="SAV753"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC68463.1"
/db_xref="GI:29604393"
/translation="MPELDRTTVALTTIGIGVPAGRDVRDPAARGYDLRLAADRA VGRREPRVVRNIPRRRGY"
3755..4408
/note="SAV754"
/codon_start=1
/transl_table=11
/product="putative TetR-family transcriptional regulator"
/protein_id="BAC68464.1"

CDS

CDS

CDS

CDS

CDS

CDS

db_xref="GI:29604394"
/translation="MTEITTSQOVGLRSEKQHEITROLIDPHATRLFLAQCFPQTTI
/AIIAAARAKTNTNIPARKEDLADHQEEFTASLANAVADNRPGESALALRAPI
BDAIQNPVAFGSGPAFAMVADSPITLACRGLNDQHEHLEAQLADATYTDPTDIT
VHIAAGLGTQTVRIVFORIQDITLGHNDRIIAHTITAEATQAFDLLEPAHAATA"
4587 . 5627
/note="SAV755"
/codon_start=1
/transl_table=11
/product="putative dehydrogenase"
/protein_id="BAC68465.1"
/db_xref="GI:29604395"
/translation="MTSPKRRITTSPPVACPATGSPSSQATSAPSYPAVTRACSTWY
CSRIORRPAAPVRPSQSDPEADFPALVSRITSDITLPGERHVLSSVILDSFDV
LYTRAGATNDAPALRIETBRPVRENPCTCGARAPAPCAHPRMAVSGRPASDDP
ATSRPAPASPNGRPNQHRGSAVSTMEAATPKEMALVLAFLDVHATGALRIHLR
AARGSVASGAGLGGDYGMAANAKGAVNLTSMALDHAPEVAVNAGHGPTLHGSG
YAAAPHRDVNAS"
complement(6101 . 8071)
/note="SAV756"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC68466.1"
/db_xref="GI:29604396"
/translation="MERSKDRRPVSLQDRAVYMPDRLTARSTGLMGIVITICLAAVV
PSIALWMTDPSPSRLSGRLAAVWRSSAETRLICTYQCPILMLSLALGVALICV
SLTVGVITTLGLAERLAELSRGSTYLBEGHAYVIGMSQVSTVIGELYAQSSTRPAA
VVLAEKRDTEHERLAALAAHVGPAIGRTVLCVSGSPSDGVIALVSPQTASTVLIPSG
EFTAAEAVTEMLALRAVALGEGTGPPVLAVALIDRRAVAPRALAGRGVTLIEDYAT
ALILAQCVRGPLSLVLDLIDFADBEHLEAFAHFGPGGALMLGHIALSCVGLLT
AEGRTLLNPATLIVPGRSLVFLTRDDGSAEPDECHRLVPSAIIAAGPEPPAHIL
ILLGNRRAPALVNNQIRRTARPGSLVDVTDRAVGPTRPESGAARLAVRFOGASL
RETLIGLDIAVPELDLITVPERGDPGRPDRLITLTLALRIEERTGEMRVITEL
TDDRRCPLAVPVRGSDVIVSGEITRGLLAQIAQNRIHAPVDELFSADGSLCLRQNG
HYTRRGCTETRTVAAARDGECHIGTRRDRAIIPDHGIRLNPDKSERREWSADD
QVVVAATPEPAFPAAGDTEHQGPADTRA"
9086 . 10072
/note="SAV757"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC68467.1"
/db_xref="GI:29604397"
/translation="MGRKRGKIGIALVTSVNLVGSASACGGSGCDKANGDDAKAGKGP
ASASPKKPAAPMLLETITPQGTGTGVAMPISVVFTEPVAKAKAAVEKMKVSA
SQPVASGVHMEGDKKADRAIDPKTVPSGTRKVIDADMGNUNGNRVEGVSHYTHFKIG
DDVAVADGVTGDKMVKVYRNGKVRVTLSTNAGSAQYPTMNGTMAVIDQEKYHNSGV
GISCEKSGPNITDITLIPMDVHLTQSGITVHYSTEDPTPGSGARSGHCVALSHSDAK
WPEYGVKKQGDPTVITIGSPRAPAPANGVADENLMDWLAGSASGETYATL"
complement(10889 . 11494)
/note="SAV758"
/codon_start=1
/transl_table=11
/product="putative acetyltransferase"
/protein_id="BAC68468.1"
/db_xref="GI:29604398"
/translation="MQRAVPRAIAMAVARPROIIECGDFVLRSWRGCRDPAPAFKLT
ESLIDLRMEPEVAVASHESTRDLAKSESKNAGDVAVYALADGTLIMCGQYRA
PEOGRIGIYMLIPATGOGIATRAVTAAMVTPLPVEYIEVHDLANTSSAIPAR
LGFTEVREDOYTPAASGSGIDVVMRLNRTTPVDSSRRP"
11580 . 12161
/note="SAV759"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC68469.1"
/db_xref="GI:29604399"
/translation="WRSRAPTRCWPSPGSTRSDATLCAWRNRRTGCTGCPRPFDRI
SAACFRVAPALLTOARPGKILITLISGWLGYARVALTVADVDGAEGLQDQVTSF
MSARSHAPVYGNPAMHATGLPEKARTTRHSPEPITATBEALHRLPAQSAVSDQM
ITVGVNHLVDVVTGSRDHPHWRGCLQCTRRGR"

CDS

db_xref="GI:29604394"
/translation="MTEITTSQOVGLRSEKQHEITROLIDPHATRLFLAQCFPQTTI
/AIIAAARAKTNTNIPARKEDLADHQEEFTASLANAVADNRPGESALALRAPI
BDAIQNPVAFGSGPAFAMVADSPITLACRGLNDQHEHLEAQLADATYTDPTDIT
VHIAAGLGTQTVRIVFORIQDITLGHNDRIIAHTITAEATQAFDLLEPAHAATA"
4587 . 5627
/note="SAV755"
/codon_start=1
/transl_table=11
/product="putative dehydrogenase"
/protein_id="BAC68465.1"
/db_xref="GI:29604395"
/translation="MTSPKRRITTSPPVACPATGSPSSQATSAPSYPAVTRACSTWY
CSRIORRPAAPVRPSQSDPEADFPALVSRITSDITLPGERHVLSSVILDSFDV
LYTRAGATNDAPALRIETBRPVRENPCTCGARAPAPCAHPRMAVSGRPASDDP
ATSRPAPASPNGRPNQHRGSAVSTMEAATPKEMALVLAFLDVHATGALRIHLR
AARGSVASGAGLGGDYGMAANAKGAVNLTSMALDHAPEVAVNAGHGPTLHGSG
YAAAPHRDVNAS"
complement(6101 . 8071)
/note="SAV756"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC68466.1"
/db_xref="GI:29604396"
/translation="MERSKDRRPVSLQDRAVYMPDRLTARSTGLMGIVITICLAAVV
PSIALWMTDPSPSRLSGRLAAVWRSSAETRLICTYQCPILMLSLALGVALICV
SLTVGVITTLGLAERLAELSRGSTYLBEGHAYVIGMSQVSTVIGELYAQSSTRPAA
VVLAEKRDTEHERLAALAAHVGPAIGRTVLCVSGSPSDGVIALVSPQTASTVLIPSG
EFTAAEAVTEMLALRAVALGEGTGPPVLAVALIDRRAVAPRALAGRGVTLIEDYAT
ALILAQCVRGPLSLVLDLIDFADBEHLEAFAHFGPGGALMLGHIALSCVGLLT
AEGRTLLNPATLIVPGRSLVFLTRDDGSAEPDECHRLVPSAIIAAGPEPPAHIL
ILLGNRRAPALVNNQIRRTARPGSLVDVTDRAVGPTRPESGAARLAVRFOGASL
RETLIGLDIAVPELDLITVPERGDPGRPDRLITLTLALRIEERTGEMRVITEL
TDDRRCPLAVPVRGSDVIVSGEITRGLLAQIAQNRIHAPVDELFSADGSLCLRQNG
HYTRRGCTETRTVAAARDGECHIGTRRDRAIIPDHGIRLNPDKSERREWSADD
QVVVAATPEPAFPAAGDTEHQGPADTRA"
9086 . 10072
/note="SAV757"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC68467.1"
/db_xref="GI:29604397"
/translation="MGRKRGKIGIALVTSVNLVGSASACGGSGCDKANGDDAKAGKGP
ASASPKKPAAPMLLETITPQGTGTGVAMPISVVFTEPVAKAKAAVEKMKVSA
SQPVASGVHMEGDKKADRAIDPKTVPSGTRKVIDADMGNUNGNRVEGVSHYTHFKIG
DDVAVADGVTGDKMVKVYRNGKVRVTLSTNAGSAQYPTMNGTMAVIDQEKYHNSGV
GISCEKSGPNITDITLIPMDVHLTQSGITVHYSTEDPTPGSGARSGHCVALSHSDAK
WPEYGVKKQGDPTVITIGSPRAPAPANGVADENLMDWLAGSASGETYATL"
complement(10889 . 11494)
/note="SAV758"
/codon_start=1
/transl_table=11
/product="putative acetyltransferase"
/protein_id="BAC68468.1"
/db_xref="GI:29604398"
/translation="MQRAVPRAIAMAVARPROIIECGDFVLRSWRGCRDPAPAFKLT
ESLIDLRMEPEVAVASHESTRDLAKSESKNAGDVAVYALADGTLIMCGQYRA
PEOGRIGIYMLIPATGOGIATRAVTAAMVTPLPVEYIEVHDLANTSSAIPAR
LGFTEVREDOYTPAASGSGIDVVMRLNRTTPVDSSRRP"
11580 . 12161
/note="SAV759"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC68469.1"
/db_xref="GI:29604399"
/translation="WRSRAPTRCWPSPGSTRSDATLCAWRNRRTGCTGCPRPFDRI
SAACFRVAPALLTOARPGKILITLISGWLGYARVALTVADVDGAEGLQDQVTSF
MSARSHAPVYGNPAMHATGLPEKARTTRHSPEPITATBEALHRLPAQSAVSDQM
ITVGVNHLVDVVTGSRDHPHWRGCLQCTRRGR"

CDS

db_xref="GI:29604394"
/translation="MTEITTSQOVGLRSEKQHEITROLIDPHATRLFLAQCFPQTTI
/AIIAAARAKTNTNIPARKEDLADHQEEFTASLANAVADNRPGESALALRAPI
BDAIQNPVAFGSGPAFAMVADSPITLACRGLNDQHEHLEAQLADATYTDPTDIT
VHIAAGLGTQTVRIVFORIQDITLGHNDRIIAHTITAEATQAFDLLEPAHAATA"
4587 . 5627
/note="SAV755"
/codon_start=1
/transl_table=11
/product="putative dehydrogenase"
/protein_id="BAC68465.1"
/db_xref="GI:29604395"
/translation="MTSPKRRITTSPPVACPATGSPSSQATSAPSYPAVTRACSTWY
CSRIORRPAAPVRPSQSDPEADFPALVSRITSDITLPGERHVLSSVILDSFDV
LYTRAGATNDAPALRIETBRPVRENPCTCGARAPAPCAHPRMAVSGRPASDDP
ATSRPAPASPNGRPNQHRGSAVSTMEAATPKEMALVLAFLDVHATGALRIHLR
AARGSVASGAGLGGDYGMAANAKGAVNLTSMALDHAPEVAVNAGHGPTLHGSG
YAAAPHRDVNAS"
complement(6101 . 8071)
/note="SAV756"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC68466.1"
/db_xref="GI:29604396"
/translation="MERSKDRRPVSLQDRAVYMPDRLTARSTGLMGIVITICLAAVV
PSIALWMTDPSPSRLSGRLAAVWRSSAETRLICTYQCPILMLSLALGVALICV
SLTVGVITTLGLAERLAELSRGSTYLBEGHAYVIGMSQVSTVIGELYAQSSTRPAA
VVLAEKRDTEHERLAALAAHVGPAIGRTVLCVSGSPSDGVIALVSPQTASTVLIPSG
EFTAAEAVTEMLALRAVALGEGTGPPVLAVALIDRRAVAPRALAGRGVTLIEDYAT
ALILAQCVRGPLSLVLDLIDFADBEHLEAFAHFGPGGALMLGHIALSCVGLLT
AEGRTLLNPATLIVPGRSLVFLTRDDGSAEPDECHRLVPSAIIAAGPEPPAHIL
ILLGNRRAPALVNNQIRRTARPGSLVDVTDRAVGPTRPESGAARLAVRFOGASL
RETLIGLDIAVPELDLITVPERGDPGRPDRLITLTLALRIEERTGEMRVITEL
TDDRRCPLAVPVRGSDVIVSGEITRGLLAQIAQNRIHAPVDELFSADGSLCLRQNG
HYTRRGCTETRTVAAARDGECHIGTRRDRAIIPDHGIRLNPDKSERREWSADD
QVVVAATPEPAFPAAGDTEHQGPADTRA"
9086 . 10072
/note="SAV757"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC68467.1"
/db_xref="GI:29604397"
/translation="MGRKRGKIGIALVTSVNLVGSASACGGSGCDKANGDDAKAGKGP
ASASPKKPAAPMLLETITPQGTGTGV

[illegible]

DB: 6 Gaps: 6

US-10-069-353A-8 (1-275) x E07846 (1-1410)

QY 7 ProthnsSergInGlnValGlyGlnMetTyrAspLeuValThrProleuLeuAsnSerVal 26

Db 841 CCTCTTCCCTCCGGAAGTGGGGGAGACTACTACGACCGTTTGACCGACTCATGATGCTGG 782

QY 27 AlaGlyGlyProCysAlaIleHisGlyTyrTrpGluAsnAspGlyValArgLaseTrp 46

Db 781 CTGGGTGGG-----AACACCCACCTCCGATCTGGCCGACCGGGGGAGCGGAGTTCA 728

QY 47 Gln----GlnAlaAlaAspArgLeuThrAspLeuValAlaGluArgThrValLeuAspGly 65

Db 727 CCGGGGCAAGCGCCGACCGCGCTCACCATCTCTCATCGGACAGCTGAGAGCATACG 668

QY 66 GlyValArgLeuLeuAspValGlyCysGlyThrGlyGlnProAlaLeuArgValAlaArg 85

Db 667 GGCCTCCCGGGTCTGTGACCTCGGCTCGGGTTCCGGAAAGCCGGCGGGTGGCTCCCGCTG 608

QY 86 AspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGlnValAlaIleAla 105

Db 607 AGCGGCGCCCTGGAGTCTGTGGCGAGCGAGGAGGAGGTTCAAGTCCGGCTGGACCC 548

QY 106 AspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSerCysValAspAlaMet 125

Db 547 GCTCTCGGGAACAGTCCGACCGCGGCGAGCCGGGTGTGTTTACCCGTCCGACCGGATG 488

QY 126 SerLeuProTyrProAspAsnAlaPheAspAlaAlaTTPAlaMetGlnSerLeuGlu 145

Db 487 GAACCTCCGGTTCCTCCGACCGGTTCCTTCGACCGGGCGTGGCCCTGAGTGTCTCTGCAC 428

QY 146 MetSerGluProAspArgAlaIleArgGluIleAspArgValLeuAspProGlyGlyIle 165

Db 427 ATGCCCAAGCCCGGACAGGTGATCCGGGAGATCGCCCGGGTCTCGCCCGGCGGG 368

QY 166 LeuGlyValThrGluValValIlyAspArgGluAlaGlyGlyMetProValSerGlyAsp 185

Db 367 CTGGCGGTCACGAGAGTGACTGCGGCGCT-TCGGGCGGACCGGACATGAAGCGGGGAGT 309

QY 186 ArgTrpProThrGlyLeuArgIle-----CysLeuAlaGluGlnLeuGluSer 202

Db 308 GCACGTCCCAAGTGTCTCGCGGTCGCCGCCGCGCTGTGTGACATGCACAGATACGCCGGAATG- 250

QY 203 LeuArgAlaAlaGlyPheGlnIleLeuAspTrpGluAspValSerSerArgThrArgTyr 222

Db 249 ATCGCGGACGCGGGCGTGGATCTGATGATGATGACGACGACATGCGGAGATCAGGTC----- 196

QY 223 PheMetProGlnPheAla-----GluGluLeuAlaAlaIleGlnHisGlyIleAla 239

Db 195 GTGGGCGCTCTTTGGCGCGCGTGGGTGACACGTAACGAGACGACCTGCAGCAGATACCG 136

QY 240 AspArgTyrGlyProAlaValAlaGlyTrp-----AlaAlaAlaValCysAspTyrGlu 257

Db 135 GCGGCGCTTCGGGATGGCGGTCCGAGATCGGGAAGCGGTGTCACAGTGCACAGCGCTC 76

QY 258 LysTyrAlaHisAspMetGlyTyrAlaIleLeuThrAlaArgLysPro 273

Db 75 CCTTGAGCGCGGACATCGGCTATGTCGTGACCGCGCGCGCGCG 28

RESUL.T.13

AX598613

LOCUS AX598613 846 bp DNA linear PAT 14-FEB-2003

DEFINITION Sequence 21 from Patent WO02088176.

ACCESSION AX598613

VERSION AX598613.1 GI:28398761

KEYWORDS

SOURCE

ORGANISM

Streptomyces platensis

Streptomyces platensis

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

Streptomycinae; Streptomycetaceae; Streptomyces.

REFERENCE

1 Farnet, C.M., Zazopoulos, E., Staiff, A. and Yang, X.

TITLE		Genes and proteins for the biosynthesis of polyketides	
JOURNAL		Patent: WO 02088176-A 21 07-NOV-2002; Ecopia Biosciences Inc. (CA)	
FEATURES		Location/Qualifiers	
source		1..846 /organism="Streptomyces platensis" /mol_type="unassigned DNA" /db_xref="taxon:58346" /note="Subsp. tobaceus"	
ORIGIN			
Alignment Scores:			
Pred. No.:	2, 21e-24	Length:	846
Score:	438.50	Matches:	106
Percent Similarity:	56.10%	Conservative:	55
Best Local Similarity:	36.93%	Mismatches:	95
Query Match:	30.37%	Indels:	31
DB:	6	Gaps:	9
US-10-069-353A-8 (1-275) x AX598613 (1-846)			
QY	2	LeuProGlyValAlaProThrSerGlnGlnValGlyGlnMetTyAspLeuValThPro	21
DB	34	GTTCCTCCCTCTCCGCCCCCGCCCGGAGAGGTGGACACTCTTCGACCGCGCTCACCGCA	93
QY	22	LeuLeuAsnSerValAlaGlyGlyProCysAlaIleHsHsGlyTYTTPGlu---Asn	40
DB	94	CTGACACCGGACCGGCGCGGCGC-----ACCTTCACCTCGGCTACTGGAGACTGCAC	147
QY	41	AspGlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg	60
DB	148	GACACACGACCGCGCTCGTGAGAGGGCGGACCGGCTCACCGACAGATACGACCGCGC	207
QY	61	ThrValLeuAspGlyValArgLeuLeuAspValGlyCysGlyThnGlyGlnProAla	80
DB	208	CTGCGGATCGACGAGGAGGAGGCTTCGACGTGCTCGAGTCTCGAGTCTGCGACCGGCGC	267
QY	81	LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGln	100
DB	268	ATGCGGATGCGCCGGGCGACCGGCGCCATGCTCAGCGGCACTGGGATCGACGACGACG	327
QY	101	ValAlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSer	120
DB	328	ATGCGCGCGCCACCGCCCTTCGCGGAGCGCGGCGCTCGACGACCGGCTGAGTTCCG	387
QY	121	CysValAspAlaMetSerLeuProTyProAspAsnAlaPheAspAlaAlaTrpAlaMet	140
DB	388	CACCGCGACCGCATTGAGACTGCTCCCTCCCGACGACTCTTCGACGCGCGCATCGCATC	447
QY	141	GlnSerLeuLeuGlnUetSerGluProAspArgAlaIleArgGluIleLeuArgValLeu	160
DB	448	GAGTCGATCTTCCACATGCGCCGACCGGAGCGGCTCTCGCGGAGATCCGCGGTACTG	507
QY	161	LyseProGlyValGlyLeuGlyValThrGluValValTySarGluAlaGlyGlyMet	180
DB	508	CGCCCGGCGCGCGCTGCTCTCACCGACTTCTTCGAGCGGCGC-----	552
QY	181	ProValSerGlyAspArgTrpProThrGly-----LeuArg	192
DB	553	CCGCTCCCGCGGAGAGGACCGCGGAGGACCGGCTCTCCGCGACTTCATCATGACG	612
QY	193	IleCysLeuAlaGluGlnLeuLeuGlnUetSerLeuArgAlaAlaGlyPhe-----Glu	209
DB	613	CTGCGCGCGCGCCGAGACTGATGCGCATGCTGCGCGGACGAGCGCTCGCTTCGTGAG	672
QY	210	IleLeuAspTrpGluAspValSerSerArgThrArgTyR---PheMetProGlnPheAla	228
DB	673	CTCTCGACATCACGAGGACGCTGCTGACCTTCGACGACGATGAGCGAGGCTCC	732
QY	229	GluGlnLeu---AlaAlaHisGlnHisGlyIleAlaAspArgTyGlyProAlaValAla	247
DB	733	CAGGAGATCGACGCGCTTTCGACGACGAGCGAGGAGAAAGTTTCAGCGCC-----	783

Qy	161	LYSPROGLYGLYILEUENGLYVALTHRGUVALVALYSAARGUVALAGLYGLYMET	180
Dy	51741	CGCCCGGGGGCGCGCTGGTCTCAACGACCTTCTCGAGCGGCG	51785
Qy	181	PROVALSERGLYASPARTRPRPROTHRGLY-----LEUDARG	192
Dy	51786	CCCGTCCCGCGCAGAAAGACGCCCGCGGTGACCGGCTCTCCGCACTTCATCATGACG	51845
Qy	193	ILECYELEUVALAGLUCHINEUENGLUSERLEUARGVALAAGLYPHE-----GLU	209
Dy	51846	GTGCGCCGCGCGCAGACTAGTACGTGCCCATGTCTGCGGACGCAAGGCTCGCTTCGTGAG	51905
Qy	210	ILELEUASPTTRPGLYUSRYVALSERSEARGTRHARGTYR---PHEMERPROGLNHEALA	228
Dy	51906	CTCTTGACATCATCCAGCAGACGCGCTTCAGACCTTGACAGATGATGACGAGGCTCC	51965
Qy	229	GLUGLUDEU---ALALAHSGLNIHGLYILEVALASPARTYGLYPROVALVALALA	247
Dy	51966	CAGAGATGCAGACCGCTTTCGACGACGAGCGCAGAGAAAGTTTCAGCCCC	52016
Qy	248	GLYTRPALLAALALALALCYASAPRTYGLYUPLYTRVALHISAPMETGLYTRVALALE	267
Dy	52017	-----GCTTCCTGATGATCGACGTTCGACGAATTC-----GCTTCGCTTCTG	52055
Qy	268	LEUTHRALAARGLYSPROVAL	274
Dy	52056	CTGACGCGCCCAAAAGCCCCCTC	52076
RESULT 15			
LOCUS	AX089434	900 bp	DNA
DEFINITION	Sequence 19 from Patent WO0116303.		linear
ACCESSION	AX089434		
VERSION	AX089434.1		GI:13443695
KEYWORDS			
SOURCE			
ORGANISM			
	Saccharopolyspora spinosa		
	Saccharopolyspora spinosa		
	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;		
	Pseudonocardineae; Pseudonocardiales; Saccharopolyspora.		
REFERENCE	1		
AUTHORS	Eberz, G., Moehrle, V., Froede, R., Velten, R. and Salas, J. A.		
TITLE	Nucleic acids which code for the enzyme activities of the spinosyn		
JOURNAL	biosynthesis		
	Patent: WO 0116303-A 19 08-MAR-2001;		
	BAYER AG (DE)		
FEATURES			
source	location/qualifiers		
	1..900		
	/organism="Saccharopolyspora spinosa"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:60894"		
	1..900		
	/note="unamed protein product; ORF7; O-Methyltransferase"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CA34880.1"		
	/db_xref="GI:13443696"		
	/translation="MNSRGANTGAPPTADQVESIFDALAHGRPLHHGYAGCYREANG		
	AAATWSDADQLTDLFTDKALIRGAGHIFDGCNGGPRVRAACASGVRVTGIVTNAOH		
	LAATRLANETGLAGSEFLVDGQOLPYDDGFPOAMQSVQIVDDAALIREVHR		
	ILBGRGFVIGDITITRVLRPEEYAAVWTGTAHTLNSFALVSEAGFEILEVDTLTAQ		
	TRGVMSYTVDEILARKIDELAGAPPAVGTIQQRVLDDIAKKGPRPAQLIAVAEYRK		
	HPDYARNESWGFMILQARKKQS"		
ORIGIN			
Alignment Scores:			
Score No.:	3,996-24	length:	900
Score:	435.50	Matches:	108
Percent Similarity:	51.67%	Conservative:	47
Best Local Similarity:	36.00%	Mismatches:	96
Query Match:	30.16%	Indels:	49
DB:	6	Gaps:	11

US-10-069-353A-8 (1-275) x AX089434 (1-900)

```
OY      7 ProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThrProLeuLeuAsnSerVal 26
      34 CCGACCGCGGATCATGGTGGATGCATCTTCGATGCGTTG-----GCG 75
OY      27 AlaGlyGlyProCysAlaIleHisGlyTyrTyrP-----GluAsnAsp 41
      76 CACGGGCGGTCCC-----CTGCACACCGGTTACTGGGCGGGGATATCGGAGAGATGCC 129
OY      42 GlyArgAlaSerTyrGlnGlnIleAlaAspArgLeuThrAspLeuValAlaGluArgThr 61
      130 GGTGCCACACCGGTGGTGGATGTCGCCACCACTGACCGACCTGTTCATCGACCAAGGCC 189
OY      62 ValLeuAspGlyGlyValArgLeuLeuAspValGlyCysGlyThrGlyGlnProAlaLeu 81
      190 GGGCTTCGTCGCCGAGCGGACGCACTGTTCCGACTGGGGTGGGCAATGGGCAGCCCTGATGC 249
OY      82 ArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGlnVal 101
      250 CGTGGCGGCATGGCCGACGCGCGTTCGACATCACCGGATCACCGTGAACGCCACGATCTC 309
OY      102 AlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspHisSerCys 121
      310 GCCGCCGCTCACCGCTCGCAACGACGACGCACTGGCCGCGCATGCTTGTGAGTTGATCTA 369
OY      122 ValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTyrAlaMetGln 141
      370 GTTCGACGGCGCCGACGCTGCTTACCCTGACGCGTTCCTTCAAGCCGCAATGGCGATGCCAG 429
OY      142 SerLeuLeuGluMetSerGlnProAspArgAlaIleArgGluIleLeuArgValLeuIys 161
      430 TTCGTCTGTGCAGATGCTGACCGACGCGCGCGATCCGCGAGGTCCACCGAATCTCGGAA 489
OY      162 ProGlyGly-----IleLeuGly-----ValThrGluValValIysArgGluAlaGly 177
      490 CCGGGGGCGCGGTGCTCTCGAGACATCATCACTCGGGTTCGACTCCCGGAAAGATAC 549
OY      178 GlyGlyMetProValSerGlyAspArgTyrProThrGlyLeuArgIleCysLeuAlaGlu 197
      550 GCGGGGGTT-----TGG---ACGGGACGACCGCCCATACCTTGAAAC 588
OY      198 GlnLeuLeuGlnSerLeuArgAlaIleGlyPheGluIleLeuAspTyrPgluAspValSer 217
      589 AGCTTCACGCGCGCTGGTTCAGCGAGCGCGGTTTCGAGATTCTCGAAGTACCGACCTCAG 648
OY      218 SerArgThrArgTyrPheMetProGlnPheAlaGluGluLeu----- 231
      649 GCACAGACCGAGGTGATGGTCTCTGTGTACGTGCAGAGAGTTGCTCCGGAAACTCGATGAG 708
OY      232 -----AlaAlaHisGlnHis-----GlyIle 238
      709 CTCGCCGGCGGTGAGCGCTGCGGCTGCGGCACCTACACGACAGGCTACTGGGAGACATC 768
OY      239 AlaAspArgTyrGlyProAlaValAlaGlyTyrAlaAlaAlaValCysAspTyrGluIys 258
      769 GCGGCGAAGCACGACGCGGACACGACAGCTGATCGCGCGGTGCGGAAATACCGGAAA 828
OY      259 -----TyrAlaHisAsp-----MetGlyTyrAlaAlaIleLeuThrAlaArgIys 272
      829 CATCGGATTCGCGCAAAACAGAAAGCAATGGATTGATGCTCTCGAGGCTCGAAG 888
```

Search completed: August 29, 2005, 21:22:47
Job time : 4877 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 29, 2005, 17:33:28 ; Search time 612 Seconds
(without alignments)
2660.014 Million cell updates/sec

Title: US-10-069-353A-8

Perfect score: 1444
Sequence: 1 VLPGAPTSQQVGMVDTLT.....YKVAHDMGYALTARKPVG 275

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-Q/cgn2.1/USPTO.spool/US1006353/rnatc.26082005.172053.14656/app.query.fasta_1.455
-DB=N Geneseq.16Dec04 -QFMT=fasta -SUFIX=mg -MINMATCH=0.1 -LOOFC=0
-LOOPEXT=0 -UNIT=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human0.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=DCE -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MAXLEN=200000000
-USER=US1006353 @CGN 1.1 470 @rnatc.26082005.172053.14656 -NCPY=6 -ICPU=3
-NO MMAP -LARGEOUTERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq.16Dec04.*

1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001a.*
5: geneseqn2001b.*
6: geneseqn2002a.*
7: geneseqn2002b.*
8: geneseqn2003a.*
9: geneseqn2003b.*
10: geneseqn2003c.*
11: geneseqn2003d.*
12: geneseqn2004a.*
13: geneseqn2004b.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1444	100.0	828	4	AAf88318
2	1444	100.0	45624	4	AAf88315
3	1444	100.0	50000	4	AAf88312
4	1444	100.0	80161	2	AA221501
5	1315	91.1	36538	10	ABV75558

Result ID	Score	Query Match	Length	DB ID	Description
6	439.5	30.4	36538	10	ABV75558
7	439	30.4	1410	2	AA080524
8	438.5	30.4	846	10	AA54227
9	438.5	30.4	52101	10	AA54217
10	438.5	30.2	900	4	AAf88324
11	435.5	30.2	45624	4	AAf88315
12	435.5	30.2	50000	4	AAf88312
13	435.5	30.2	80161	2	AA221501
14	379	26.1	6085	3	AA770153
15	376.5	26.1	852	2	AA55802
16	376.5	26.1	852	10	ADE10221
17	376.5	26.1	53500	3	AA55842
18	376.5	26.1	53500	10	ADE10261
19	375.5	26.0	25681	10	ADJ53165
20	361.5	25.0	103599	4	ABX04971
21	345	23.9	828	3	AA55801
22	345	23.9	828	10	ADE10220
23	343.5	23.8	12381	3	AA55831
24	320.5	22.2	2122	2	AA770152
25	268	18.6	52659	9	AA160188
26	260	17.3	843	8	AB280088
27	250	17.3	843	10	ADD19104
28	247.5	17.3	858	13	ADT44187
29	247.5	17.1	7205	13	ADR84263
30	246.5	17.1	843	8	AB280087
31	246.5	17.1	843	10	ADD19103
32	242.5	16.8	1444	6	AA231672
33	241.5	16.7	1444	6	AB553042
34	240	16.6	930	8	AB280091
35	238.5	16.5	1383	8	AA292614
36	232	16.1	1044	8	AB280079
37	232	16.1	1044	8	ADJ78872
38	229.5	15.9	691	10	ADK54920
39	229.5	15.9	1038	8	AB280075
40	229.5	15.9	1038	10	ADD19093
41	229.5	15.9	1038	12	ADQ95177
42	229.5	15.9	1038	12	ADQ95182
43	229.5	15.9	8179	12	ADQ95181
44	229.5	15.9	8179	12	AB280080
45	228	15.8	1044	8	AB280080

ALIGNMENTS

RESULT 1
ID AAF88318 standard; DNA; 828 BP.
XX
AC AAF88318;
XX
DT 28-AUG-2001 (first entry)
XX
DE S. spinosa DNA fragment encoding ORF1, SEQ ID 7.
XX
KW Forosamine; trimethylurhamose; polyketide synthase; biosynthesis;
KW spinosyn; polyketide aglycone; transgenic plant; insect resistance;
KW meclofen; insecticide; O-methyltransferase; dr.
XX
OS Saccharopolyspora spinosa.
XX
PN DE19957268-A1.
XX
PD 08-MAR-2001.
XX
PF 29-NOV-1999; 99DE-01057268.
XX
PR 27-AUG-1999; 99DE-01040596.
XX
PA (FARB) BAYER AG.
XX
PI Eberz G, Moehle V, Froede R, Velten R, Salas JA;
DR WPI; 2001-267102/28.

DR P-PSDB; AAB70948.

XX New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for
PT recombinant production of insecticidal spinosyns and their derivatives.

XX Claim 7a; Page 102-104; 354pp; German.

XX This invention describes a novel method nucleic acid (I) and its encoded
CC polypeptide (II) containing at least one region that encodes an enzymatic
CC activity involved in biosynthesis of spinosyns. (I) are used (i) to
CC identify, inactive or modulate genes involved in the biosynthesis of
CC (ii); (ii) to generate a library of polypeptide synthases; (iii) for
CC adding forosamine or trimethylrhinamose to a spinosyn or polypeptide
CC aglycone; and (iv) for recombinant production of the corresponding
CC enzymes, which are used for production of (ii), their precursors or
CC derivatives, including production of transgenic plants that express (ii)
CC and thus have increased resistance to insects. (I) are also useful as
CC markers for sequencing of the Saccharopolyspora spinosa genome. (II) are
CC macrolides with insecticidal, but not antibacterial, activity, and can
CC also be used to raise specific antibodies, useful for identifying
CC expression clones in a gene bank. Cells transformed with (I) may produce
CC (II) at significantly increased levels or produce new derivatives of
CC (II). This sequence, ORF 1, encodes an S. spinosa O-methyltransferase
CC protein

XX SQ Sequence 828 BP; 140 A; 262 C; 285 G; 141 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8 36e-133 Length: 828
Score: 1444.00 Matches: 275
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatch: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-069-353a-8 (1-275) x AAF88318 (1-828)

QY 1 ValLeuProGlyValAlaProThrsGergInGlnValGlyGlnMetTyAspLeuValThr 20
DB 1 GTGTTGCCAGGTGGCGCACCAATCGACGAGTTGGGCAATGATGACCTGTCTACG 60
QY 21 ProLeuLeuAsnSerValAlaGlyProCySAAlaIleHisGlyTyrTrpGluAn 40
DB 61 CGTTGCTGAATCTCGTCCGGGCGGCCCTCGCATCCACACGGCTACCTGAGGAGAC 120
QY 41 AspGlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
DB 121 GACGGCGGGCTTCTGCGACAGAGCGCCGACCGGCTCACCGACTTGTCCGAAACGG 180
QY 61 ThrValLeuAspGlyValAlaArgLeuLeuAspValGlyCySAIleThrGlyInProAla 80
DB 181 ACCGTCGTGATGAGCGCGCTTCGACTGCTCGATGAGGAGTGGGACCGACCAACGAG 240
QY 81 LeuArgValAlaAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGln 100
DB 241 CTGGCGCTGCGCGGCGACCAACCGATCCAGATCACCGGATCACCGTCAAGCGAGTGCMA 300
QY 101 ValAlaIleAlaAlaAspCySAIleArgGluArgGlyLeuSerHisArgValAspPheSer 120
DB 301 GTGGCATGTCGGCTGATTTGGCCAGCGACCGAGCTAAGCCACCGGCTGACTTCTCG 360
QY 121 CysValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
DB 361 TGGCTGATGATGCTCTGCTGCGTACCGGACATGCTTGAAGCGCGCTGAGGACATG 420
QY 141 GlnSerLeuLeuGlnMetSerGluProAspArgAlaIleArgGluIleLeuArgValLeu 160
DB 421 CAGTGGCTGTGGAGATGTCCAAACCGACCGTGCATCCGGGAATCTTCTGATACTC 480
QY 161 LysProGlyGlyIleLeuGlyValThrGluValValIleArgGluAlaGlyValGlyMet 180
DB 481 AAACCCGGTGCATCTCTGGGCTCACCGAGTGTCTCAACAGAGAGCGGCGGCGGATG 540

QY 181 ProValSerGlyAspArgTrpProThrGlyLeuArgIleCySAleuAlaGluGlnLeu 200
DB 541 CCGGTGTCGGGAGACAGTGGCCGACCGGCTTGAGATCTCGCTGGCTGACCACTTCTG 600
QY 201 GluSerLeuArgAlaAlaGlyPheGlnIleLeuAspTrpGluAspValSerSerArgThr 220
DB 601 GAATGCTGCGTGCAGCGGGGTTGAGATCTCGATTGGAGGACGTCTCTCGAGGACC 660
QY 221 ArgTyrPheMetProGlnPheAlaGluGlnLeuAlaIleHisGlnIleGlyTleAlaAsp 240
DB 661 CGGACTTCAAGCCGCACTTGGCCGAGAGCTTCCTGGCCACCGACGGAATCCGGAC 720
QY 241 ArgTyrGlyProAlaValAlaGlyTyrAlaAlaAlaValCySAAspTyrGluLysTyrAla 260
DB 721 AGTACGGCGCGGCTGTGCGCGGCTGGGCGCCGCGGCTCGATTATGAGAAATATGCC 780
QY 261 HisAspMetGlyTyrAlaIleLeuThrAlaArgLysProValGly 275
DB 781 CACGACATGGGCTATGCAATTCTGACCGCGCGGAAAGCGGTCCGC 825

RESULT 2

AAF88315/C
ID AAF88315 standard; DNA; 45624 BP.

AAF88315;

28-AUG-2001 (first entry)

S. spinosa DNA fragment SEQ ID 4.

Forosamine; trimethylrhinamose; polypeptide synthase; biosynthesis;

spinosyn; polypeptide aglycone; transgenic plant; insect resistance;

macrolide; insecticidal; de.

Saccharopolyspora spinosa.

DE19957268-A1.

08-MAR-2001.

29-NOV-1999; 99DE-01057268.

27-AUG-1999; 99DE-01040596.

(FARB) BAYER AG.

Eberz G, Moehrl V, Froede R, Velten R, Salas JA;

WPI; 2001-267102/28.

Claim 7; Page 58-74; 354pp; German.

This invention describes a novel method nucleic acid (I) and its encoded

polypeptide (II) containing at least one region that encodes an enzymatic

activity involved in biosynthesis of spinosyns. (I) are used (i) to

identify, inactive or modulate genes involved in the biosynthesis of

(ii); (ii) to generate a library of polypeptide synthases; (iii) for

adding forosamine or trimethylrhinamose to a spinosyn or polypeptide

aglycone; and (iv) for recombinant production of the corresponding

enzymes, which are used for production of (ii), their precursors or

derivatives, including production of transgenic plants that express (ii)

and thus have increased resistance to insects. (I) are also useful as

markers for sequencing of the Saccharopolyspora spinosa genome. (II) are

macrolides with insecticidal, but not antibacterial, activity, and can

also be used to raise specific antibodies, useful for identifying

expression clones in a gene bank. Cells transformed with (I) may produce

(II) at significantly increased levels or produce new derivatives of

(II). This sequence represents a genomic DNA fragment of the S. spinosa

genome which contains the coding regions for proteins involved in

forosamine and trimethylrhinamose biosynthesis

```

XX      SQ      Sequence 45624 BP, 7933 A; 14913 C; 14940 G; 7839 T; 0 U; 0 Other:
Alignment Scores:
Pred. No.:      1,26e-130      Length:      45624
Score:          1444.00      Matches:      275
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:            4      Gaps:      0
US-10-069-353A-8 (1-275) x AAF88315 (1-45624)

QY      1 ValLeuProGlyGlyAlaProThrSerGlnGlnValGlyGlnMetYrAspLeuValThr 20
      828 GTGTTCACAGTGGCCACACCATCGACAGAGTTGGGAGATGATGATGACCTGTGTCACG 769
QY      21 ProLeuLeuAsnSerValAlaGlyGlyProCysAlaIleHisIleGlyYrTrpGluAsn 40
      768 CCGTGTCTGAACCTCGGTCCGGGGCCCTCGGCCATCCACACCGGCTACTGGAGAAC 709
QY      41 AspGlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
      708 GACGGGGGGGCTTCCTGGACGACGCGCCGACCGGCTCACGACCTTGTGCGCGAAACGG 649
QY      61 ThrValLeuAspGlyGlyValArgLeuLeuAspValGlyCysGlyYrThrGlyGlnProAla 80
      648 ACCGTGCTCATGGCGCGCTTCGATCGATGTGGGGTGCGATGCCGGAACACACGCG 589
QY      81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyYrIleThrValSerGlnValGln 100
      588 CTGCGCGCTCCGGCGACACACCGGATCCAGATCACCGGCTACCGTCCAGCCAGGTCCAA 529
QY      101 ValAlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120
      528 GTGGCCATCCCGCTGATTCGGACCGGAAACGGGATTAAGCCACCGGGTGACTTCTCG 469
QY      121 CysValAspAlaMetSerLeuProYrProAspAsnAlaPheAspAlaAlaTPAlaMet 140
      468 TGGGTGATGCCATGTCTCCGCTACCCGACCAATCTTTCGACCGCGCTGGCCATG 409
QY      141 GluSerLeuLeuGluMetSerGluProAspArgAlaIleArgGluIleLeuArgValLeu 160
      408 CAGTCCGTGTGGAGATGCCAACCGGACCGGTCCATCCGGGAATCTTCCAGTATCTC 349
QY      161 LysProGlyGlyIleLeuGlyValThrGluValAlaLysArgGluAlaGlyGlyMet 180
      348 AAACCCGGTGGCATCTCCGCGCTCACCGAGGTGTCAAAAGAGAGCGGGCGGGGATG 289
QY      181 ProValSerGlyAspArgTrpProThrGlyLeuArgIleCysLeuAlaGluGlnLeu 200
      288 CCGGTGTCCGGGGGACAGGTGGCCGACCGGCTTCGAGTCTGCGGTGAGCACTTCTG 229
QY      201 GluSerLeuArgAlaAlaGlyPheGluIleLeuAspTrpGluAspValSerSerArgThr 220
      228 GAATCGCTGGTGAGCGGGGTTCGAGATCTCATTTGGAGGAGCGTGTCTCAGGAC 169
QY      221 ArgYrPheMetProGlnPheAlaGluIleuAlaAlaHisIleGlnIleGlyIleAlaAsp 240
      168 CGGTACTTCATGCGGACGTTCCGGAAGAGTCTCGTCCGACACGACGCGGATGGGGAC 109
QY      241 ArgYrGlyProAlaValAlaGlyYrTrpAlaAlaAlaValCysAspYrGlyLysYrTrpAla 260
      108 AGGTACCGGGCGCGCTGTCCGCGGTGGCGCCCGCGGTCTGCGATTTAGAAATATGACC 49
QY      261 HisAspMetGlyYrTrpAlaIleLeuThrAlaArgLysProValGly 275
      48 CACGACATGGGCTATGGATTCTAAGCGCGGAGCGGATCGGC 4

```

```

AC      AAF88312;
XX      DT      28-AUG-2001 (first entry)
XX      DE      S. spinosa DNA fragment SEQ ID 1.
XX      KW      Forosamine; trimethylrhannose; polyketide synthase; biosynthesis;
XX      KW      spinosyn; polyketide aglycone; transgenic plant; insect resistance;
XX      KW      macrolide; insecticidal; ds.
XX      OS      Saccharopolyspora spinosa.
XX      PN      DE19957268-A1.
XX      PD      08-MAR-2001.
XX      PF      29-NOV-1999; 99DE-01057268.
XX      PR      27-AUG-1999; 99DE-01040596.
XX      PA      (PARB ) BAYER AG.
XX      PI      Eberz G, Moehle V, Froede R, Velten R, Salas JA;
XX      DR      WPI; 2001-267102/28.
XX      PT      New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for
XX      PS      recombinant production of insecticidal spinosyns and their derivatives.
XX      PS      Claim 7; Page 14-31; 354pp; German.
XX      CC      This invention describes a novel method nucleic acid (I) and its encoded
XX      CC      polypeptide (II) containing at least one region that encodes an enzymatic
XX      CC      activity involved in biosynthesis of spinosyns. (I) are used (i) to
XX      CC      identify, inactivate or modulate genes involved in the biosynthesis of
XX      CC      (II); (ii) to generate a library of polyketide synthases; (iii) for
XX      CC      adding forosamine or trimethylrhannose to a spinosyn or polyketide
XX      CC      aglycone; and (iv) for recombinant production of the corresponding
XX      CC      enzymes, which are used for production of (II), their precursors or
XX      CC      derivatives, including production of transgenic plants that express (II)
XX      CC      and thus have increased resistance to insects. (I) are also useful as
XX      CC      markers for sequencing of the Saccharopolyspora spinosa genome. (II) are
XX      CC      macrolides with insecticidal, but not antibacterial, activity, and can
XX      CC      also be used to raise specific antibodies, useful for identifying
XX      CC      expression clones in a gene bank. Cells transformed with (I) may produce
XX      CC      (II) at significantly increased levels or produce new derivatives of
XX      CC      (II). This sequence represents a genomic DNA fragment of the S. spinosa
XX      CC      genome which contains the coding regions for proteins involved in
XX      CC      forosamine, trimethylrhannose and polyketide synthase biosynthesis
XX      SQ      Sequence 50000 BP, 8484 A; 16290 C; 16535 G; 8691 T; 0 U; 0 Other:
Alignment Scores:
Pred. No.:      1.41e-130      Length:      50000
Score:          1444.00      Matches:      275
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:            4      Gaps:      0
US-10-069-353A-8 (1-275) x AAF88312 (1-50000)

QY      1 ValLeuProGlyGlyAlaProThrSerGlnGlnValGlyGlnMetYrAspLeuValThr 20
      44797 GTGTTCACAGTGGCCACACCATCGACAGAGTTGGGAGATGATGATGACCTGTGTCACG 44856
QY      21 ProLeuLeuAsnSerValAlaGlyGlyProCysAlaIleHisIleGlyYrTrpGluAsn 40
      44857 CCGTGTCTGAACCTCGGTCCGGGGCCCTCGGCCATCCACACCGGCTACTGGAGAAC 44916
QY      41 AspGlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
      44917 GACGGGGGGGCTTCCTGGACGACGCGCCGACCGGCTCACGACCTTGTGCGCGAAACGG 44976

```

[illegible]

```
FT misc_feature 24102..25349
FT /tag= u
FT /note= "Beta-ketosynthase domain (KS1) : part of extender
FT module 1"
FT 25683..26684
FT /tag= v
FT /note= "Acyl transferase domain (AT1) : part of extender
FT module 1"
FT 27582..28121
FT /tag= w
FT /note= "Beta-ketoreductase domain (KR1) : part of extender
FT module 1"
FT 28404..28649
FT /tag= x
FT /note= "Acyl carrier protein domain (ACP1) : part of
FT extender module 1"
FT 29024..30295
FT /tag= z
FT /note= "Beta-ketosynthase domain (KS2) : part of extender
FT module 2"
FT 30629..31621
FT /tag= aa
FT /note= "Acyl transferase domain (AT2) : part of extender
FT module 2"
FT 31697..32254
FT /tag= ab
FT /note= "Dehydratase domain (DH2) : part of extender module
FT 2"
FT 33035..34072
FT /tag= ac
FT /note= "Enoyl reductase domain (ER2) : part of extender
FT module 2"
FT 34082..34621
FT /tag= ad
FT /note= "Beta-ketoreductase domain (KR2) : part of extender
FT module 2"
FT 34886..350295
FT /tag= ae
FT /note= "Acyl carrier protein domain (ACP2) : part of
FT extender module 2"
FT 35419..44931
FT /tag= af
FT /product= "spnc"
FT /note= "Spinosyn biosynthesis protein C"
FT 35518..36786
FT /tag= ag
FT /note= "Beta-ketosynthase domain (KS3) : part of extender
FT module 3"
FT 37108..38097
FT /tag= ah
FT /note= "Acyl transferase domain (AT3) : part of extender
FT module 3"
FT 38916..35374
FT /tag= ai
FT /product= "spnc"
FT /note= "Spinosyn biosynthesis protein C"
FT 38992..39528
FT /tag= aj
FT /note= "Beta-ketoreductase domain (KR3) : part of extender
FT module 3"
FT 39790..40035
FT /tag= ak
FT /note= "Acyl carrier protein domain (ACP3) : part of
FT extender module 3"
FT 40102..41373
FT /tag= al
FT /note= "Beta-ketosynthase domain (KS4) : part of extender
FT module 4"
FT 41713..42705
FT /tag= am
FT /note= "Acyl transferase domain (AT4) : part of extender
FT module 4"
FT 43615..44157
FT misc_feature
```

```
FT /tag= am
FT /note= "Beta-ketoreductase domain (KR4) : part of extender
FT module 4"
FT 44431..44676
FT /tag= an
FT /note= "Acyl carrier protein domain (ACP4) : part of
FT extender module 4"
FT 44966..59752
FT /tag= ao
FT /product= "spnd"
FT /note= "Spinosyn biosynthesis protein D"
FT 45077..46348
FT /tag= ap
FT /note= "Beta-ketosynthase domain (KS5) : part of extender
FT module 5"
FT 46691..47674
FT /tag= aq
FT /note= "Acyl transferase domain (AT5) : part of extender
FT module 5"
FT 47753..48310
FT /tag= ar
FT /note= "Dehydratase domain (DH5) : part of extender module
FT 5"
FT 49226..49771
FT /tag= as
FT /note= "Beta-ketoreductase domain (KR5) : part of extender
FT module 5"
FT 50009..50254
FT misc_feature

Alignment Scores:
Pred. No.: 2,55e-130 Length: 80161
Score: 1444.00 Matches: 275
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-069-353a-8 (1-275) x AA221501 (1-80161)

QY 1 ValLeuProGlyValAlaProThSerGInGInValGInMeTyRAspLeuValThr 20
DB GTGTGCGAGTGGCGCACCAACATCGACAGTGGGAGATGATGACCTGTGCACG 20227

QY 21 ProLeuLeuAsnSerValAlaGlyIProCVAlaIleHisIaGlyTYrTrpGluAsn 40
DB CCGTTGCTGAACCTGGTGGGCGGCGCCCTGTGCCATTCACACGACTGAGAGAAC 20287

QY 41 AspGlyARGLaseTrpGInGInAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
DB 20288 GACGGCGGGCTTCCTGCGACGACGCCGCGACCGGCTACCGACTGTGTCCGACCGG 20347

QY 61 ThrValLeuAspGlyValAlaArgLeuLeuAspValGlyCVAsGlyThrGlyInProAla 80
DB 20348 ACCGTGCTGATGGCGGCGTTGCACTGCTGATGTGGGGTGGCGGTACCGGACAAACGCG 20407

QY 81 LeuArgValAlaArgAspAsnAlaIleGInIleThrGlyTleThrValSerGInValGln 100
DB 20408 CTGGCGGTCCGGCGCAACGCGATCCCAATCAACCGGCACTCCGTCAGCCGACAGTGCAC 20467

QY 101 ValAlaIleAlaAlaAspCVAsAlaArgGluArgGlyLeuSerHisArgValaAspPheSer 120
DB 20468 GTGGCCATCGCCCTGATTCGCGACGCGAACGCGACTAACGACGCGGGTGCATCTTCG 20527

QY 121 CysValaAspAlaMetSerLeuProTYrProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
DB 20528 TGCCTGATCCATGTCCTCGCGGTACCGCGCAATGCTTTCACGCCCGCTGGCCATG 20587

QY 141 GluSerLeuLeuGluMetSerGluProAspArgAlaIleArgGluIleLeuArgValaLeu 160
DB 20588 CAGTCGCTGTTGAGATGTCGAACCGGACCGGATCCGGAATCTTTCAGATGATCTC 20647

QY 161 LysProGlyGlyIleLeuGlyValThrGluValaValaLysArgGluAlaGlyGlyMet 180
DB |||||
```

Db 20648 AAACCGGATCCTCGGCGTCAACGAGTCGTCAAAAGAAAGCGGCGCGGATG 20707
 Qy 181 ProValSerGlyAspArgTrpProThrGlyLeuArgIleCysLeuAlaGluGlnLeu 200
 Db 20708 CCGGCTCCGGGACAGGTGGCCGACCGGCTTCGGATTCGCTGAGCACTTCTG 20767
 Qy 201 GluSerLeuArgAlaAlaGlyPheGluIleLeuAspTrpGluAspValSerSerArgThr 220
 Db 20768 GAATCGCTCGTGCAGCGGGGTTCCGAGATCTCGATTGGAGAGACGTCGTGAGAGACC 20827
 Qy 221 ArgTyrPheMetProGlnPheAlaGluGluLeuAlaAlaHisGlyIleAlaAsp 240
 Db 20828 CCGTACTTCATGCGGAGTTCGCCGGAAGAGCTCGCTGCGCACGACGCGGATCGCGGAC 20887
 Qy 241 ArgTyrGlyProAlaValAlaGlyTyrAlaAlaAlaValAlaCysAspTyrGluIleTyrAla 260
 Db 20888 AGGTACGGGCGCGGCTGTCCCGCTGGCGCCGCCGCTGTGCGATTATGAGAAATATGCC 20947
 Qy 261 HisAspMetGlyTyrAlaIleLeuThrAlaArgIleAspProValGly 275
 Db 20948 CACGACATGGCTATGCGATCTCGACGCGCGGAAGCCGCTCGGC 20992
 RESULT 5
 ABV75558/c
 ID ABV75558 standard; DNA; 36538 BP.
 XX
 XX ABV75558;
 DT 22-JAN-2003 (first entry)
 XX
 XX Saccharopolyspora butenyl-spinosyn biosynthetic gene cluster 2.
 XX
 XX Butenyl; biosynthetic enzyme; PKS; polyketide synthetase; macrolide;
 KW metabolite; spinosyn; gene; ds.
 XX
 OS Saccharopolyspora sp.
 XX
 XX Key Location/Qualifiers
 FH CDS Complement(114..338)
 FT /tag= a
 FT /product= "busF"
 FT /note= "No start codon given"
 FT 1389..2561
 FT /tag= b
 FT /product= "busG"
 FT 2601..3353
 FT /tag= c
 FT /product= "busH"
 FT complement(3359..4546)
 FT /tag= d
 FT /product= "busI"
 FT 4684..6303
 FT /tag= e
 FT /product= "busJ"
 FT 6317..7510
 FT /tag= f
 FT /product= "busK"
 FT 7555..8406
 FT /tag= g
 FT /product= "busL"
 FT /note= "No start codon given"
 FT 8640..9572
 FT /tag= h
 FT /product= "busM"
 FT /note= "No start codon given"
 FT complement(9668..10666)
 FT /tag= i
 FT /product= "busN"
 FT complement(10675..12135)
 FT /tag= j
 FT /product= "busO"
 FT complement(12864..14177)
 FT /tag= k
 FT CDS

FT /product= "busP"
 FT 14627..15970
 FT /tag= l
 FT /product= "busQ"
 FT /note= "No start codon given"
 FT 16008..17144
 FT /tag= m
 FT /product= "busR"
 FT /note= "No start codon given"
 FT 17168..17917
 FT /tag= n
 FT /product= "busS"
 FT complement(18520..19932)
 FT /tag= o
 FT /product= "ORF LI"
 FT complement(19978..20488)
 FT /tag= p
 FT /product= "ORF LII"
 FT /note= "No start codon given"
 FT complement(20536..21033)
 FT /tag= q
 FT /product= "ORF LIII"
 FT /note= "No start codon given"
 FT 21179..21925
 FT /tag= r
 FT /product= "ORF LIV"
 FT complement(22671..23453)
 FT /tag= s
 FT /product= "ORF LVI"
 FT complement(23687..24886)
 FT /tag= t
 FT /product= "ORF LVII"
 FT complement(26177..26923)
 FT /tag= u
 FT /product= "ORF LVIII"
 FT /note= "No start codon given"
 FT 27646..28476
 FT /tag= v
 FT /product= "ORF LIX"
 FT /note= "No start codon given"
 FT
 PN WO200279477-A2.
 XX
 XX 10-OCT-2002.
 PD
 XX 28-MAR-2002; 2002WO-US009968.
 XX
 PF 30-MAR-2001; 2001US-0280175P.
 XX
 PK (DOWC) DOW AGROSCIENCES LLC.
 XX
 PA Hahn DR, Jackson JD, Bullard BS, Gustafson GD, Waldron C;
 PI Mitchell JC;
 PI
 XX WPI; 2003-058434/05.
 DR P-PSDB; ABP57683, ABP57684, ABP57685, ABP57686, ABP57687, ABP57688,
 DR ABP57689, ABP57690, ABP57691, ABP57692, ABP57693, ABP57694, ABP57695,
 DR ABP57696, ABP57697, ABP57698, ABP57699, ABP57700, ABP57701, ABP57702,
 DR ABP57703, ABP57704.
 XX
 XX New butenyl-spinosyn biosynthetic genes, useful for increasing the
 FT production of butenyl-spinosyn insecticidal macrolides, or for changing
 FT the metabolites or products produced by spinosyn-producing
 FT microorganisms.
 PT
 XX
 PS Claim 2; Page 99-119; 218pp; English.
 XX
 XX The invention relates to a novel DNA molecule comprising a DNA sequence
 CC that encodes a butenyl-spinosyn biosynthetic enzyme, a butenyl-spinosyn
 CC PKS (polyketide synthetase) domain, or a spinosyn PKS module. The butenyl-
 CC -spinosyn biosynthetic genes are useful for increasing the production of
 CC butenyl-spinosyn insecticidal macrolides. The genes are also useful for
 CC changing the metabolites or products produced by spinosyn-producing

CC microorganisms. The present sequence represents a DNA molecule encoding
 CC butenyl-epinosyn biosynthetic enzymes

XX Sequence 36538 BP; 6867 A; 12266 C; 11182 G; 6223 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4 93e-118	Length:	36538
Score:	1315.00	Matches:	252
Percent Similarity:	94.18%	Conservative:	7
Best Local Similarity:	91.64%	Mismatches:	16
Query Match:	91.07%	Indels:	0
DB:	10	Gaps:	0

US-10-069-353a-8 (1-275) x ABV75558 (1-36538)

```

OY 1 ValLeuProGlyGlyValAlaProThSerGlnGlnValGlyGlnMetTyrAspLeuValThr 20
   |||||
DB 938 GTGTTGCCAGGTGGCGCTGCAACATCGACGAGTGGGCAATGATGACCTGTGACCC 879
OY 21 ProLeuLeuAsnSerValAlaGlyGlyProGlyAlaIleHisHisGlyTyrTrpGluAsn 40
   |||||
DB 878 CCGTGTCTGAATTCGGTCCGCGCGCGCCCTGCGCATCCACACGCTACTGGAGAAC 819
OY 41 ABPGlyArGAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
   |||||
DB 818 GACGGGGGACCTTCTGGGACGAGCGCGCGACCGGCTCAACGATCTTGTCCGCGAACG 759
OY 61 ThrValLeuAspGlyGlyValArgLeuLeuAspValGlyCysGlyTyrGlyGlnProAla 80
   |||||
DB 758 ACCGCGCTGATGGCGGCAATGACCTGCGGATGTGGGGTGGCGGACCGGACCAACGAG 699
OY 81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyTyrIleThrValSerGlnValGln 100
   |||||
DB 698 CTGCGCGTGGCGCGGCAACGCGATCCGATCCGCGCATCCGCGCATCCGCGCATCCGCG 639
OY 101 ValAlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120
   |||||
DB 638 GCTGCGATGCGCTTATTTGGTGGCGCGGAAACGAGACTAAGCCACGAGTGGACTTCTG 579
OY 121 CysValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
   |||||
DB 578 TGGTGCACCCCATGTCCTCCGCGGATCCGCGCAATGCTTCAACGCTGCTGGGCGATC 519
OY 141 GlnSerLeuLeuGlnUetSerGluProAspArgAlaIleArgGluIleLeuArgValLeu 160
   |||||
DB 518 CAGTCCGCTGTGAGATGTGGAACCGGACCGTGCATCCGGGAAATCGTTCAAGTGTCTC 459
OY 161 LysProGlyGlyIleLeuGlyValThrGluValValLysArgGluAlaGlyGlyMet 180
   |||||
DB 458 AAACCCGGTGGCATCTCCGAGTCAACGAGTGTCAAAAGAGAAATAGGACAGCGGATC 399
OY 181 ProValSerGlyAspArgTrpProThrGlyLeuArgIleCysLeuAlaGlnLeu 200
   |||||
DB 398 CCGGTGTCTGGGACATGTGGCGACCGGCTCCGATCTGCTGGCTGAGCACTTCTG 339
OY 201 GlnSerLeuArgAlaAlaGlyPheGluIleLeuAspTrpGluAspValSerSerArgThr 220
   |||||
DB 338 GAATCGCTGTGTCAGCGGGTTCGAAATCTCCCTGTGACGACGCTGTGTCGCGGACC 279
OY 221 ArgTyrPheMetProGlnPheAlaGlnGlnLeuAlaAlaHisGlnHisGlyTyrIleAlaAsp 240
   |||||
DB 278 CGGACTTCAATGCCGCGATGCCGAGCGCTCCCTCGGACCGGACCGGAGATCGCGAA 219
OY 241 ArgTyrGlyProAlaValAlaGlyTyrAlaAlaAlaValCysAspTyrGluLysTyrAla 260
   |||||
DB 218 AGGTACGGGGCGGTGTGGCGCATGGGGCGCGCGCTGCGATTAAGAAATACGCC 159
OY 261 HisAspMetGlyTyrAlaIleLeuThrAlaArgLysProValGly 275
   |||||
DB 158 GACGACATGGGCTATGCGATTTTGAACGCGCGGAAACCGGCTGGCG 114

```

RESULT 6
 ABV75558

ID	ABV75558 standard; DNA; 36538 BP.
XX	ABV75558;
AC	22-JAN-2003 (first entry)
XX	Saccharopolyspora butenyl-epinosyn biosynthetic gene cluster 2.
DT	Butenyl; biosynthetic enzyme; PKS; polyketide synthetase; macroide;
XX	metabolite; epinosyn gene; ds.
XX	Saccharopolyspora sp.
OS	Location/Qualifiers
XX	complement(114..938)
FH	/*tag= a
FT	/product= "bueF"
FT	/note= "No start codon given"
FT	1389..2561
FT	/*tag= b
FT	/product= "bueG"
FT	2601..3353
FT	/*tag= c
FT	/product= "bueH"
FT	complement(3359..4546)
FT	/*tag= d
FT	/product= "bueI"
FT	4684..6303
FT	/*tag= e
FT	/product= "bueJ"
FT	6317..7510
FT	/*tag= f
FT	/product= "bueK"
FT	7555..8406
FT	/*tag= g
FT	/product= "bueL"
FT	/note= "No start codon given"
FT	8640..9572
FT	/*tag= h
FT	/product= "bueM"
FT	/note= "No start codon given"
FT	complement(9668..10666)
FT	/*tag= i
FT	/product= "bueN"
FT	complement(10675..12135)
FT	/*tag= j
FT	/product= "bueO"
FT	complement(12864..14177)
FT	/*tag= k
FT	/product= "bueP"
FT	14627..15970
FT	/*tag= l
FT	/product= "bueQ"
FT	/note= "No start codon given"
FT	15008..17144
FT	/*tag= m
FT	/product= "bueR"
FT	/note= "No start codon given"
FT	17168..17917
FT	/*tag= n
FT	/product= "bueS"
FT	complement(18520..19932)
FT	/*tag= o
FT	/product= "ORF LI"
FT	complement(19978..20488)
FT	/*tag= p
FT	/product= "ORF LI"
FT	/note= "No start codon given"
FT	complement(20536..21033)
FT	/*tag= q
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= r
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= s
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= t
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= u
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= v
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= w
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= x
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= y
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= z
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AA
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AB
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AC
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AD
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AE
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AF
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AG
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AH
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AI
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AJ
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AK
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AL
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AM
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AN
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AO
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AP
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AQ
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AR
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AS
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AT
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AU
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AV
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AW
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AX
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AY
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= AZ
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BA
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BB
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BC
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BD
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BE
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BF
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BG
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BH
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BI
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BJ
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BK
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BL
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BM
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BN
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BO
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BP
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BQ
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BR
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BS
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BT
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BU
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BV
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BW
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BX
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BY
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= BZ
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= CA
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= CB
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= CC
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= CD
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= CE
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= CF
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= CG
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= CH
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= CI
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= CJ
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= CK
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= CL
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= CM
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= CN
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= CO
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"
FT	21179..21925
FT	/*tag= CP
FT	/product= "ORF LIIT"
FT	/note= "No start codon given"


```
FT FT /*tag= r
FT FT /product= "ORF LIV"
FT CDS complement(22671..23453)
FT FT /*tag= s
FT FT /product= "ORF LVT"
FT CDS complement(23687..24886)
FT FT /*tag= t
FT FT /product= "ORF LVII"
FT CDS complement(26177..26923)
FT FT /*tag= u
FT FT /product= "ORF LVIII"
FT FT /note= "No start codon given"
FT CDS 27646..28476
FT FT /*tag= v
FT FT /product= "ORF LIX"
FT FT /note= "No start codon given"
XX PN WO200279477-A2.
XX PD 10-OCT-2002.
XX PF 28-MAR-2002; 2002WO-US009968.
XX PR 30-MAR-2001; 2001US-0280175P.
XX PA (DOWC ) DOW AGROSCIENCES LLC.
XX PI Hahn DR, Jackson JD, Bullard BS, Gustafson GD, Waldron C;
PI Mitchell JC;
XX WI: 2003-058434/05.
XX DR P-PDB; ABP57683, ABP57684, ABP57685, ABP57686, ABP57687, ABP57688,
XX DR ABP57689, ABP57690, ABP57691, ABP57692, ABP57693, ABP57694, ABP57695,
XX DR ABP57696, ABP57697, ABP57698, ABP57699, ABP57700, ABP57701, ABP57702,
XX DR ABP57703, ABP57704.
XX PT New butenyl-spinosyn biosynthetic genes, useful for increasing the
XX PT production of butenyl-spinosyn insecticidal macrolides, or for changing
XX PT the metabolites or products produced by spinosyn-producing
XX PT microorganisms.
XX PS Claim 2; Page 99-119; 218pp; English.
XX CC The invention relates to a novel DNA molecule comprising a DNA sequence
XX CC that encodes a butenyl-spinosyn biosynthetic enzyme, a butenyl-spinosyn
XX CC PKS (polyketide synthetase) domain, or a spinosyn PKS module. The butenyl-
XX CC spinosyn biosynthetic genes are useful for increasing the production of
XX CC butenyl-spinosyn insecticidal macrolides. The genes are also useful for
XX CC changing the metabolites or products produced by spinosyn-producing
XX CC microorganisms. The present sequence represents a DNA molecule encoding
XX CC butenyl-spinosyn biosynthetic enzymes
XX SQ Sequence 36538 BP; 6867 A; 12266 C; 11182 G; 6223 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9,42e-32 Length: 36538
Score: 439.50 Matches: 105
Percent Similarity: 52.36% Conservative: 50
Best Local Similarity: 35.47% Mismatches: 100
Query Match: 30.44% Indels: 41
DB: 10 Gaps: 8

US-10-069-353A-8 (1-275) x ABV75558 (1-36538)
OY 7 ProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThrProLeuLeuAsnSerVal 26
DB 7540 CCAAGCGGAGTACAGTGAAGTCCATCTTCGAC-----GCCGTG 7578
OY 27 AlaGlyGlyProCysAlaIleHisHisGlyTyrTrp-----GluAsnAsp 41
DB 7579 GCGCAGAGGCGT---GCCCTGCACCAACGATTACTGGCGCGCGGCGGATTCGGAGGATGCC 7635
OY 42 GlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArgThr 61
```

```
DB 7636 GGGGCCACACCTTGGTGGAGCGCTCCGACCACTCGACGACGACTTTCATGCAGCAAGGCC 7695
OY 62 ValLeuAspGlyGlyValArgLeuAspValGlyCysGlyThrGlnProAlaLeu 81
DB 7696 GGGCTTCCCGCCCGGAGCCACCTGTTCCAGCTGGGCTTGGCAATGGGACACCCCTAGTCT 7755
OY 82 ArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGlnVal 101
DB 7756 CGCGCGGACGACCAAGAGCGTTGAGTACCGGAAATCACCGTGAACGCCGACATCTC 7815
OY 102 AlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSerCys 121
DB 7816 GCCCGCGTACACGAGCTCGCCAAACGACGAGCTGGCGCAGACGTCCTGGTTGCATCTA 7875
OY 122 ValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaIleThrAlaMetCln 141
DB 7876 GTCGACGCGCGCGGCTCCCTTACCCGGAAGGTTCTTTCACGCGCGCATGGGCGATGCA 7935
OY 142 SerLeuLeuGluMetSerGluProAspArgAlaIleArgIleLeuArgValLeuAsp 161
DB 7936 TCCGTGTACAGATGTCGACCAAGCTCCCGATCCGCGAGTCCACCGAATCTTGAA 7995
OY 162 ProGlyGlyIleLeuGlyValThrGluValValysArgGluAlaGlyGlyMetPro 181
DB 7996 CCGGCGGCGCAGTTCGTCCTCGGGGAGATCATCACTGCT-----GCTCGACTCCG 8046
OY 182 ValSerGlyAspArgTrpProThrGlyLeuArgIleCysIleAlaGluGlnLeuGlu 201
DB 8047 GAAAGTACCGCGGGGTGGACCGCGCACGCCCTTGAACGCTTGAACGCTTACCCGCG 8106
OY 202 SerLeuArgAlaAlaGlyPheGluIleLeuAspTrpGluAspValSerSerArgThrArg 221
DB 8107 CTGTAACCGAAGCGCGGTTTGAATTTCTCAAGTCAACCGACCTTACGCGCGCACACGA 8166
OY 222 TyrPheMetProGlnIlePheAlaGluGluLeu----- 231
DB 8167 TGCATGTGCTCTGTATGTGACGAGATGTTGCTCCGGAACTCGATGAGCTCGCGGCTC 8226
OY 232 -----AlaAlaHisGlnHis-----GlyIleAlaAspArgTyr 242
DB 8227 GAGCTCGGCTGTGCGGACCTTCCACGACGACGACTTGGAGACATTCGCGGAGAC 8286
OY 243 GlyProAlaValAlaGlyTyrPalaAlaAlaValCysAspTyrGluLys-----Tyr 259
DB 8287 GACCCGGAGACAGCGAGCTGATCGCCGCTCCGGAATACCGGAACATCCCGATTA 8346
OY 260 AlaHisAsp-----MetGlyTyrAlaIleLeuThrAlaArgLys 272
DB 8347 GCCAGAAACGAGAAAGCATGGTTCATGCTCTGACGAGGCGGAAAG 8394

RESULT 7
AA080524/C
ID AA080524 standard; DNA; 1410 BP.
XX AC
XX AA080524;
XX AC
XX 10-APR-1995 (first entry)
XX DT
XX Oxidoreducing avermectin DNA from S. avermectilis ATCC31271.
XX DE
XX Modified; oxidoreducing; avermectin; S. avermectilis; wild type; disabled;
XX KW
XX Streptomyces avermectilis.
XX OS
XX Key Location/Qualifiers
FH misc_feature 627..1075
FT /*tag= a
FT /note= "Avermectin coding sequence"
XX FT
XX PN JP06189774-A.
```

PD 12-UTL-1994.
 XX 24-DEC-1992; 92JP-00344605.
 XX 24-DEC-1992; 92JP-00344605.
 PR (KITA) KITASATO KENKYUSHO SH.
 XX WPI; 1994-259554/32.
 XX A new modified DNA fragment encoding an avermectin deriv. - useful for
 PT the prodn. of the deriv.
 XX
 XX Claim 2; Page 2; 17pp; Japanese.
 XX
 CC The sequences in AA080524 and AA070315-16 represent modified
 CC oxidoreducing avermectin genes from *S. avermitilis* strains deposited as
 CC ATCC 31271, FERM P-13154 and FERM P-13155. These genes were introduced
 CC into *S. avermitilis* strains in which the wild type avermectin gene has
 CC been disabled such that the modified avermectins are produced in large
 CC amounts
 XX
 XX Sequence 1410 BP; 232 A; 488 C; 463 G; 227 T; 0 U; 0 Other;
 SO
 Alignment Scores:
 Pred. No.: 1.8e-33 Length: 1410
 Score: 439.00 Matches: 111
 Percent Similarity: 56.52% Conservative: 45
 Best Local Similarity: 40.22% Mismatches: 107
 Query Match: 30.40% Indels: 15
 DB: 2 Gaps: 6
 US-10-069-353A-8 (1-275) x AA080524 (1-1410)
 QY 7 ProThSerGlnGlnValGlyGlnMetCysAspLeuValThrProLeuLeuAanSerVal 26
 DB 841 CCTCTTCCCTGGAGAGGGGAGACTACTACGACCGTTGACCGACTTACGAAATCGCGC 782
 QY 27 AlaGlyGlyProCysValAlaIleHisGlyTyrTrpGluHisAspGlyArgAlaSerTrp 46
 DB 781 CTGGGTGG-----AACACCCACCTCGGATCTCGCGCGCGCGCGGAGAGGATGCA 728
 QY 47 Gln--GlnAlaAlaAspArgLeuThrAspLeuValAlaGluArgThrValLeuAspGly 65
 DB 727 CCGGCGAAGGCG 668
 QY 66 GlyValArgLeuLeuAspValGlyCysGlyThrGlyGlnProAlaLeuArgValAlaArg 85
 DB 667 GCG 608
 QY 86 AspAsnAlaIleGlnIleThrGlyTyrThrValSerGlnValAlaIleAlaAla 105
 DB 607 AGCG 548
 QY 106 AspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSerCysValAspAlaMet 125
 DB 547 GCTCTCGCGAAGCG 488
 QY 126 SerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMetGlnSerLeuLeuGlu 145
 DB 487 GAACTGCGCGTTCCCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 428
 QY 146 MetSerGluProAspArgAlaIleArgGluIleLeuArgValLeuLysProGlyGlyIle 165
 DB 427 ATGCCGACCG 368
 QY 166 LeuGlyValThrGluValValValArgGluAlaGlyGlyMetProValSerGlyAsp 185
 DB 367 CTGGCGCGTCAACG 309
 QY 186 ArgTrpProThrGlyLeuArgGlyIle-----CysLeuAlaGluGlnLeuGluSer 202
 DB 308 GCACGTCCCAAGTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 250

QY 203 LeuArgAlaAlaGlyPheGlnIleLeuAspTrpGluAspValSerSerArgThrArgTyr 222
 DB 249 ATGCCGACCG 196
 QY 223 PheMetProGlnPheAla-----GluGluLeuAlaAlaHisGlnHisGlyIleAla 239
 DB 195 GTCCG 136
 QY 240 AspArgTyrGlyProAlaValAlaGlyTrp-----AlaAlaAlaValCysAspTyrGlu 257
 DB 135 GCG 76
 QY 258 LysTyrAlaHisAspMetGlyTyrAlaIleLeuThrAlaArgLysPro 273
 DB 75 CCTGGACCG 28
 RESULT 8
 ID AAD54227
 ID AAD54227 standard; DNA; 846 BP.
 AC AAD54227;
 XX
 DT 17-JUN-2003 (first entry)
 XX
 DE Streptomyces platensis subspecies rosaceus dorrigocin ORF10 DNA.
 XX
 KW Polyketide biosynthesis; dorrigocin; DORR; lactimidomycin; LACT; gene;
 KM de.
 XX
 OS Streptomyces platensis.
 XX
 FH Key Location/Qualifiers
 FT CDS 163..1149
 FT /tag= a
 FT /product= "ORF10 protein"
 FT
 PN WO200288176-A2.
 XX
 PD 07-NOV-2002.
 XX
 PF 26-APR-2002; 2002WO-CA000591.
 XX
 PR 26-APR-2001; 2001US-0286346P.
 XX
 PA (ECOP-) ECOPIA BIOSCIENCES INC.
 XX
 PI Farnet CM, Zazopoulos E, Staffa A, Yang X;
 DR WPI; 2003-201222/19.
 DR P-PSDB; AAB35493.
 XX
 PT Novel isolated or purified polypeptide involved in biosynthesis of
 PT polyketide dorrigocin or polyketide lactimidomycin, useful for preparing
 PT dorrigocin or lactimidomycin.
 XX
 PS Claim 6; Page 198; 312pp; English.
 XX
 CC The invention relates to novel proteins involved in the biosynthesis of
 CC polyketide dorrigocin (DORR) or lactimidomycin (LACT) biosynthesis by
 CC microorganisms. Sequences of the invention allow direct manipulation of
 CC dorrigocin, lactimidomycin and related chemical structures via chemical
 CC engineering of the enzymes involved in the biosynthesis of dorrigocin and
 CC lactimidomycin. They are useful for introducing chemical handles into
 CC normally inert positions that permit subsequent chemical modifications
 CC and facilitate the development of polyketides. The genes and proteins of
 CC the invention can also be used to generate a focused library of analogues
 CC around a polyketide lead candidate to fine-tune the compound for optimal
 CC properties. They are useful for generating antibodies specific for the
 CC polyketide biosynthesis. The present sequence is Streptomyces platensis
 CC subspecies rosaceus DORR ORF10 DNA
 XX
 SO Sequence 846 BP; 141 A; 343 C; 249 G; 113 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,06e-33	Length:	846
Score:	438.50	Matches:	106
Percent Similarity:	56.10%	Conservative:	55
Best Local Similarity:	36.93%	Mismatches:	95
Query Match:	30.37%	Indels:	31
DB:	10	Gaps:	9

US-10-069-353a-8 (1-275) x AAD54227 (1-846)

```

QY      2 LeuProGlyGlyValAlaProThrSerGlnGlnValGlyGlnMetTyraSpleuValThrPro 21
      34 GTCCCGGTCCTCCGCCCCCGTCCCGAAGAGTCGGACACTCTACGACCGGCTCACCGGCA 93
QY      22 LeuLeuAsnSerValAlaGlyGlyProGlyAlaIleHisGlyTyrThrGlu---Asn 40
      94 CTGACACACCGAAGCGCGCGCGGCGC-----AGCTTCACCTCGGCTACTGGAGACGTGCAC 147
QY      41 AspGlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
      148 GACACAGACACCGCCGCTCGTGGAGAGGCGCGACCGGCTACCGACACGATGACCGACCGC 207
QY      61 ThrValLeuAspGlyGlyValArgLeuLeuAspValGlyCysGlyThrGlyGlnProAla 80
      208 CTGCGGATCGACACGAGGACAGCGGGCTCTCGACGTCGCGCTCGGAGTCGCGCAGCGCGCC 267
QY      81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGln 100
      268 ATGCGGATGCGCGCGCGGCGACCGCGCGCCCATGTCACGCGGATCGCATGCAAGACACG 327
QY      101 ValAlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120
      328 ATCGCGCGCGCGACCGGCTCGCCGACGAGGCGCGCGCTGAGCGACCGCGTGAAGTTCCG 387
QY      121 CysValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
      388 CACGCGCGACGCGCATGAGCTGCCCTTCCCGACGACTCTTGACGCGCGCATGCGCATC 447
QY      141 GlnSerLeuLeuGlnMetSerLeuProAspArgAlaIleArgGlnIleLeuArgValLeu 160
      448 GAGTGATGTTCCACATGCGCGCGACCGGCGGAGTCTCGCGGAGATCGCGCGCTACTCG 507
QY      161 LysProGlyGlyIleLeuGlyValThrGluValIleLysArgGluAlaGlyGlyGlyMet 180
      508 CGCCCGCGGCGCGCGCTGCTCTCAACGACTCTTCGAGCGCGC-----LeuArg 552
QY      181 ProValSerGlyAspArgTrpProThrGly-----LeuArg 192
      553 CCGGTCGCCGCGGACGAGCGCGGTCGACCGGCTCTCCGCACTTCATCATGACG 612
QY      193 IleCysLeuAlaGlnIleLeuGlnLeuSerLeuArgAlaIleArgPhe-----Glu 209
      613 CTGCGCGCGCGCGGACGAGCTACGTCGCTGCGCGACGACGCTCGCGCTCGTCAAG 672
QY      210 IleLeuAspTrpGluAspValSerSerArgTrpArgTyr---PheMetProGlnPheAla 228
      673 CTCCTCGACATGACGAGGAGCGTGGCTGACGACTTCGACGACATGACGAGGCGCTCC 732
QY      229 GluGluLeu---AlaAlaHisGlnHisGlyIleAlaAspArgTyrGlyProAlaValAla 247
      733 CAGGAGATGACGACCGCTTCGACGACGACGAGCGAGAAAGTTCAAGCCCC-----783
QY      248 GlyTrpAlaAlaAlaValCysAspTrpGluLysTyrAlaHisAspMetGlyTyrAlaIle 267
      784 -----GCTTCATGATGACCTCGACGCAATTC-----GCTCGGTTCTG 822
QY      268 LeuThrAlaArgLysProVal 274
      823 CTGACCGCGCAAAAGCCCTC 843

```

RESULT 9
AAD54217

```

ID      AAD54217 standard; DNA; 52101 BP.
XX
AC      AAD54217;
XX
DT      17-JUN-2003 (first entry)
XX
DE      Streptomyces platensis subspecies rosaceus dorrigocin DNA.
XX
KM      Polyketide biosynthesis; dorrigocin; DORR; laccimidomycin; LACT; gene;
XX      ds.
XX
XX      Streptomyces platensis.
XX
OS      Location/Qualifiers
XX      FH      complement(67..3720)
XX      FT      /tag= a
XX      FT      /product= "ORF1 protein"
XX      FT      /tag= b
XX      FT      /product= "ORF2 protein"
XX      FT      /note= "No start codon"
XX      FT      /partial
XX      FT      /tag= c
XX      FT      /product= "ORF3 protein"
XX      FT      /tag= d
XX      FT      /product= "ORF4 protein"
XX      FT      /tag= e
XX      FT      /product= "ORF5 protein"
XX      FT      /tag= f
XX      FT      /product= "ORF6 protein"
XX      FT      /note= "No start codon"
XX      FT      /partial
XX      FT      /tag= g
XX      FT      /product= "ORF7 protein"
XX      FT      /note= "No start codon"
XX      FT      /partial
XX      FT      /tag= h
XX      FT      /product= "ORF8 protein"
XX      FT      /tag= i
XX      FT      /product= "ORF9 protein"
XX      FT      /tag= j
XX      FT      /product= "ORF10 protein"
XX
XX      MO200288176-A2.
XX
XX      PD      07-NOV-2002.
XX
XX      PF      26-APR-2002; 2002MO-CA000591.
XX
XX      PR      26-APR-2001; 2001US-0286346P.
XX
XX      PA      (ECOP-) ECOP1A BIOSCIENCES INC.
XX
XX      PI      Farnet CM, Zazopoulos E, Staffa A, Yang X;
XX
XX      DR      WPI, 2003-201222/19.
XX      DR      P-PSDB; AAE35484, AAE35485, AAE35486, AAE35487, AAE35488, AAE35489,
XX      DR      AAE35490, AAE35491, AAE35492, AAE35493.
XX
XX      Novel isolated or purified polypeptide involved in biosynthesis of
XX      polyketide dorrigocin or polyketide laccimidomycin, useful for preparing
XX      dorrigocin or laccimidomycin.
XX
XX      Claim 1; Page 85-113; 312pp; English.
XX

```



```

Best Local Similarity: 36.00%  Mismatches: 96
Query Match: 30.16%  Indels: 49
DB: 4  Gaps: 11
US-10-069-353a-8 (1-275) x AAF88315 (1-900)

QY 7 ProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThrProLeuLeuAsnSerVal 26
DB 34 CCGACCGCGGATCAGGTGAGTCCATCTTCGATGCGTTG-----GCG 75
QY 27 AlaGlyGlyProCysAlaIleHisGlyTyrTrp-----GluAsnAsp 41
DB 76 CACGGAGCGTCC-----CTGCACCACTGTTACTGGCGCGGCGGATTCGGAGATGCC 129
QY 42 GlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArgThr 61
DB 130 GGTGGCACACCGCTGCGATGCTGCGACCACTGACCGACCTGTTCATCGACAAGGCC 189
QY 62 ValLeuAspGlyGlyValArgLeuLeuAspValGlyCysGlyThrGlyGlnProAlaLeu 81
DB 190 GCGCTCCGTCGCGGAGCGCAGCTGTTCGACTGGGCTGGCGCAATGGGAGCCCGTAGTC 249
QY 82 ArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGlnVal 101
DB 250 CGTGCGCGATGCGCGACGCGCGCTTGAGTACCGGAAATCACCGTGAACGCCAGCATCTC 309
QY 102 AlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSerCys 121
DB 310 GCCGCGCGCACCGAGCTGCGCCACAGACCGGACTGCGCGGAGCTTGAGTTGATCTA 369
QY 122 ValAlaAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaIleTrpAlaMetGln 141
DB 370 GTGCGCGCGCGCCACTGCTCCCTACCGCGGCGGCTTTCTTCAGGCGCGCATGGCGATGAG 429
QY 142 SerLeuLeuGluMetSerGluProAspArgAlaIleArgGlnIleLeuArgValLeuAsn 161
DB 430 TCCGCTGTCGATGCTGAGCCAGGCCCGCGGATTCGCGAGGTCCACCGAAATCTTGAA 489
QY 162 ProGlyGly-----IleLeuGly-----ValThrGluValValGlyArgGlnAlaGly 177
DB 490 CCGCGCGCGCGGTGCTGCTCGGAGACATCATCTCGGTTGCACTCCCGAAGACATGAC 549
QY 178 GlyGlyMetProValSerGlyAspArgTrpProThrGlyLeuArgIleCysLeuAlaGlu 197
DB 550 GCGGCGGCTT-----TGG--ACGGGACACGACCGCCCATCTTGAAC 588
QY 198 GlnLeuLeuGlnSerLeuArgAlaAlaGlyPheGlnIleLeuAspTrpGlnAspValSer 217
DB 589 AGCTTCACGCGCGCTGCTGACGAGACCGCGGTTGCGAATTTCTGAAATCACCGACCTCAG 648
QY 218 SerArgThrArgTyrPheMetProGlnPheAlaGlnGluLeu----- 231
DB 649 GCACAGACACAGGTGATGCTGCTCGTACGTCGACGAGTTGCTCCGGAACCTGATGAG 708
QY 232 -----AlaAlaHisGlnHis-----GlyIle 238
DB 709 CTCGCGCGCGCTGACCTGCGGCTGTCGACACTGACGACAACTGCTGGAGACATC 768
QY 239 AlaAspArgTyrGlyProAlaValAlaGlyTrpAlaAlaAlaValCysAspTrpGluAsn 258
DB 769 GCGGCGAGACCGCGGACGACACGACGATGCGCGGCTGGCGGAAATTCGGGAAA 828
QY 259 -----TyrAlaHisAsp-----MetGlyTyrAlaIleLeuThrAlaArgAsn 272
DB 829 CATCGGATTCGCGGAAACGAAAGCATGGGTTTCAATGCTCTCCAGGCTCGAAG 888

```

```

XX S. spinosa DNA fragment SEQ ID 4.
DE Forsamine; trimethylrhinamose; polyketide synthase; biosynthesis;
XX spinosyn; polyketide aglycone; transgenic plant; insect resistance;
XX macrolide; insecticidal; de.
OS Saccharopolyspora spinosa.
XX DE19957268-A1.
XX PD 08-MAR-2001.
XX PF 29-NOV-1999; 99DE-01057268.
XX PR 27-AUG-1999; 99DE-01040596.
XX (FARB ) BAYER AG.
XX Eberz G, Moehrle V, Froede R, Velten R, Salas JA;
XX MPI, 2001-267102/28.
DR
XX This invention describes a novel method nucleic acid (I) and its encoded
XX polypeptide (II) containing at least one region that encodes an enzymatic
XX activity involved in biosynthesis of spinosyns. (I) are used (i) to
XX identify, inactive or modulate genes involved in the biosynthesis of
XX (II); (ii) to generate a library of polyketide synthases; (iii) for
XX adding forosamine or trimethylrhinamose to a spinosyn or polyketide
XX aglycone; and (iv) for recombinant production of the corresponding
XX enzymes, which are used for production of (II), their precursors or
XX derivatives, including production of transgenic plants that express (II)
XX and thus have increased resistance to insects. (i) are also useful as
XX markers for sequencing of the Saccharopolyspora spinosa genome. (ii) are
XX macrolides with insecticidal, but not antibacterial, activity, and can
XX also be used to raise specific antibodies, useful for identifying
XX expression clones in a gene bank. Cells transformed with (I) may produce
XX (II). This sequence represents a genomic DNA fragment of the S. spinosa
XX genome which contains the coding regions for proteins involved in
XX forosamine and trimethylrhinamose biosynthesis
SQ Sequence 45624 BP, 7933 A, 14913 C, 14940 G, 7838 T, 0 U, 0 Other;
Alignment Scores:
Pred. No.: 3,08e-31 Length: 45624
Score: 435.50 Matches: 108
Percent Similarity: 51.67% Conservative: 47
Best Local Similarity: 36.00% Mismatches: 96
Query Match: 30.16% Indels: 11
DB: 4 Gaps: 49
US-10-069-353a-8 (1-275) x AAF88315 (1-45624)

QY 7 ProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThrProLeuLeuAsnSerVal 26
DB 7434 CCGACCGCGGATCAGGTGAGTCCATCTTCGATGCGTTG-----GCG 7475
QY 27 AlaGlyGlyProCysAlaIleHisGlyTyrTrp-----GluAsnAsp 41
DB 7476 CACGGAGCGTCC-----CTGCACCACTGTTACTGGCGCGGCGGATTCGGAGATGCC 7529
QY 42 GlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArgThr 61
DB 7530 GGTGGCACACCGCTGCGATGCTGCGACCACTGACCGACCTGTTCATCGACAAGGCC 7589
QY 62 ValLeuAspGlyGlyValArgLeuLeuAspValGlyCysGlyThrGlyGlnProAlaLeu 81

```

Db	7590	GGCGTCCTCCCGGAGCGCACTGTGGACCTGGGCGTGGCGCAATTGGGACCGCGTAGTC	7649
Oy	82	ArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGlnVal	101
Db	7650	CGTCCGCATCGCGACCGCGCGGTTCAGTCAACCGGAATCAACCGTGAACCGCCACGATCTC	7709
Oy	102	AlaIleAlaAlaAspCysAlaArgGlnArgGlyIleuSerHisArgValAspPheSerCys	121
Db	7710	GGCCCGCCGACCAAGCTCTGCCAAGCGAGCCGACGCGCGGCACTTGAAGTGGATCTTA	7769
Oy	122	ValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTPrAlaMetGln	141
Db	7770	GTCACGCGCGCCCGACAGCTCCCTCAACCGGACGGTTCTTTCAGGCGCATGGCGATGAG	7829
Oy	142	SerLeuLeuGlnMetSerGlnProAspArgAlaIleArgGlnIleLeuArgValLeuLys	161
Db	7830	TCCGTCGTGACAGATGTGTGACCAAGCCGCGGATCCGCGAGTCCACCGAATCCTGGAA	7889
Oy	162	ProGlyGly-----IleLeuGly-----ValThrGlnValValLysArgGlnAlaGly	177
Db	7890	CCCGGCGCGCGGATTCGTCTCTGGAGACATCATCACTCGGGTTGCACTCCCGAAGAGTAC	7949
Oy	178	GlyGlyMetProValSerGlyAspArgTrpProThrGlyLeuArgIleCysLeuAlaGln	197
Db	7950	GGCGCGGTT-----TGG--ACGGACAGACCGCCCATACCTTGAAC	7988
Oy	198	GlnLeuLeuGlnSerLeuArgAlaAlaGlyPheGlnIleLeuAspTrpGlnAspValSer	217
Db	7989	AGCTTACAGCGCGCTGTGACGGAAGCCGGGTTCCGAAATTCGAAAGTCAACGACCTCAG	8048
Oy	218	SerArgThrArgTrpPheMetProGlnPheAlaGlnGlnIleu-----	231
Db	8049	GCACAGACCAAGTGCATGATCTCTCTGTACGTCAACGAGTTGCTCCGAAACTCGATGAG	8108
Oy	232	-----AlaAlaHisGlnHis-----GlyIle	238
Db	8109	CTCCGCGCGCGTCTGAGCTCTGGCTGTCCGCACTTACAGCAACGCTACTTGGAGACATC	8168
Oy	239	AlaAspArgTrpGlyProAlaValAlaGlyTrpAlaAlaAlaValCysAspTrpGlnLys	258
Db	8169	GCGCGGAAGACCGAGCCGGACACAGACAGCTGATCGCGGGTTCGGAAATACCGGAAA	8228
Oy	259	-----TyrAlaHisAsp-----MetGlyTyrAlaIleLeuThrAlaArgLys	272
Db	8229	CATCCGAGTATCGCCAGAAACGAGGAAGCATGGGTTTCATCTCTCAGGCTCGAAG	8288
RESULT 12			
AAf88312/c			
ID	AAf88312 standard; DNA; 50000 BP.		
XX	AAf88312;		
AC			
XX	28-AUG-2001 (first entry)		
XX			
DE	S. spinosa DNA fragment SEQ ID 1.		
XX			
KW	Forosamine; trimethylrharnnose; polyketide synthase; biosynthesis;		
KM	spinosyn; polyketide aglycone; transgenic plant; insect resistance;		
KM	macrolide; insecticidal; ds.		
OS	Saccharopolyspora spinosa.		
PN	DEL19957268-A1.		
PD	08-MAR-2001.		
XX			
PF	29-NOV-1999; 99DE-01057268.		
XX			
PR	27-AUG-1999; 99DE-01040596.		
PA	(FARB) BAYER AG.		
XX			
PI	Eberz G, Moehrl V, Froede R, Velten R, Salas JA;		

XX WP1; 2001-267102/28.

XX New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for

PT recombinant production of insecticidal spinosyn and their derivatives.

XX

XX

PS- Claim 7, Page 14-31; 354pp; German.

XX

CC This invention describes a novel method nucleic acid (I) and its encoded

CC polypeptide (II) containing at least one region that encodes an enzymatic

CC activity involved in biosynthesis of spinosyns. (I) are used (i) to

CC identify, inactivate or modulate genes involved in the biosynthesis of

CC (II); (ii) to generate a library of polypeptide synthases; (iii) for

CC adding forosamine or trimethylthamose to a spinosyn or polypeptide

CC aglycone; and (iv) for recombinant production of the corresponding

CC derivatives, which are used for production of (II), their precursors or

CC derivatives, including production of transgenic plants that express (II)

CC and thus have increased resistance to insects. (I) are also useful as

CC markers for sequencing of the Saccharopolyspora spinosa genome. (II) are

CC macrocides with insecticidal, but not antibacterial, activity, and can

CC also be used to raise specific antibodies, useful for identifying

CC expression clones in a gene bank. Cells transformed with (I) may produce

CC (II) at significantly increased levels or produce new derivatives of

CC (II). This sequence represents a genomic DNA fragment of the *S. spinosa*

CC genome which contains the coding regions for proteins involved in

CC forosamine, trimethylthamose and polyketide synthase biosynthesis

XX

XX

Sequence 50000 BP; 8484 A; 16290 C; 16535 G; 8691 T; 0 U; 0 Other;

50

Alignment Scores:

Pred. No.: 3,46e-31 Length: 50000

Score: 435.50 Matches: 108

Percent Similarity: 51.67% Conservative: 47

Best Local Similarity: 36.00% Mismatches: 96

Query Match: 30.16% Indels: 49

DB: 4 Gaps: 11

US-10-069-353A-8 (1-275) x AAP88312 (1-50000)

QY 7 ProTherSergInglInValIyGlnMetTyRAspLeuValThrProLeuLeuAsnSerVal 26

DB 38191 CCGACCGCGGATCAGGAGTGAGTCCATCTTCGATGCGGTG-----GCG 38150

QY 27 AlaGlyGlyProCysAlaIleHisIleGlyTyTrp-----GluAsnAsp 41

DB 38149 CACGGGGGTGCC-----CTGCACCAACGGTTACTGTGGCGCGCGGTATCGGAGGATGCC 38096

QY 42 GlyArgIAserTrpGlnGlnAlaIAspArgLeuThrAspLeuValAgluArgThr 61

DB 38095 GGTCCCAACCGGTGTGTGATGCTGCGGACCAATCAACGACTGTTCATCCACAAGGCC 38036

QY 62 ValLeuAspGlyGlyValArgLeuLeuAspValGlyCysGlyThrGlyGlnProAlaLeu 81

DB 38035 GCGGTCTGTCGCGGAGGACGACTGTTCGACCTGGGCTGGCGGCAATGGGACGCCGATGATC 37976

QY 82 ArgValAlaArgAspAlaIleGlnIleThrGlyTleThrValSergInValGlnVal 101

DB 37975 CGTCCGCGCATCCGCGACGCGCGCTTCGAGTCAACCGGAATCAACCGTGAAGCCCGACGATTC 37916

QY 102 AlaIleAlaIAspArgAlaArgGlyArgGlyLeuSeriIAspArgValAspPheSerCys 121

DB 37915 GCCGCGCGCACCAAGCTTCGCGCAACGACACCGGACTGGCGGCGAGTTCGATTCATCA 37856

QY 122 ValAspAlaMetSerLeuProTyRProAspAlaIAspAspAlaIAspAlaIAspAlaMetGln 141

DB 37855 GTCCAGCGCGCGCCAGCGTCCGCTTACCGGAGCGTTCTTCAGCGCCGATGCGGATGACG 37796

QY 142 SerLeuLeuGlnMetSergIuPProAspArgAlaIleArgGluIleLeuArgValLeuLys 161

DB 37795 TCCGTCGTGCAGATCGTGACACGAGCGCGCGCATCCGCGAGGTCCACCGAATCCCTGGAA 37736

QY 162 ProGlyGly-----IleLeuGly-----ValThrGluValValysArgGluAlaGly 177

```
Db      37735 CCCGGCGCGGTCCTCCTCGAGACATCATCTCGGGTTCGACTCCCGGAGAGTAC 37676
Qy      178 G1YGIYMeTProVAlSerGIYAsPaRGTPrProThnGIYLeuAArgIleCYsleuAlaGIu 197
Db      37675 GCGGCGGTT-----TGG---ACGGGACAGACCGCCCATCTTGGAAC 37637
Qy      198 G1nLeuLeuGIuSerLeuAArgAlaAglYpHeGIuIleLeuAsPTpGIuSpVaISeR 217
Db      37636 AGCTTACGCGGCGCTGATCGAGACCGGGGTTCGAGATTCTCGAAGTCAACGACCTCAGC 37577
Qy      218 SerAgtThAArgTYrPhMeTProGIuPhnAlaGIuGIuLeu----- 231
Db      37576 GCACAGACCAAGTGATGGTCTCTGTACGTGACGTGCGAGTTCCTCGGAACTCGATGAG 37517
Qy      232 -----AlaAlaHisGlnHis-----G1YIle 238
Db      37516 CTCGCGGCGTGCAGGCTCGGCTGTGCGGACCTTACAGCAACGCTACTTGGAAGACATC 37457
Qy      239 AlAAsPaRgTYrGIYProAlaVaAlaAglYTPAlaAlaAlaVaICyAsAsPTYrGIuYs 258
Db      37456 GCGGCGAGACAGACGACCGGACACAGCAGTGAATCGCGCGGTTGCGGAATACCGGAA 37397
Qy      259 -----TYrAlaHisAsp-----MetGIYTYrAlaIleLeuThrAlaArgIys 272
Db      37396 CATCGGATTACGCCAGAAACAGAGAAAGCATGGTTTCATCTCTCTGACGCTCGAAAG 37337

RESULT 13
AA221501/C
ID      AA221501 standard; DNA; 80161 BP.
XX
AC      AA221501;
XX
DT      01-DEC-1999 (first entry)
XX
DE      DNA fragment of Saccharopolyspora spinosa containing biosynthetic genes.
XX
KW      Spinosyn biosynthetic enzyme; open reading frame; ORF;
KW      insecticidal microtides; arachnid; nematode; insect; polyketide;
KW      polyketide synthase; PKS; extender module; initiator module;
KW      acyl transferase domain; AT; acyl carrier protein; ACP;
KW      beta-ketosynthase domain; KS; KR; dehydratase domain; DH;
KW      enoyl reductase domain; BR; beta-ketoreductase; insecticide; ss.
OS      Saccharopolyspora spinosa.
XX
XX
Key     location/Qualifiers
FH      complement(1135..1971)
FT      /tag= a
FT      /product= "ORF16"
FT      /note= "Protein involved in transcription control"
FT      2024..2791
FT      /tag= b
FT      /product= "ORF15"
FT      /note= "keto acyl reductase"
FT      complement(3416..4165)
FT      /tag= c
FT      /product= "spnS"
FT      /note= "Spinosyn biosynthesis protein S"
FT      complement(4168..5325)
FT      /tag= d
FT      /product= "spnR"
FT      /note= "Spinosyn biosynthesis protein R"
FT      complement(5363..6751)
FT      /tag= e
FT      /product= "spnQ"
FT      /note= "Spinosyn biosynthesis protein Q"
FT      7083..8450
FT      /tag= f
FT      /product= "spnP"
FT      /note= "Spinosyn biosynthesis protein P"
FT      8967..10427
FT      /tag= g
FT      /product= "spnO"
FT
```

```
FT      /note= "Spinosyn biosynthesis protein O"
FT      10436..11434
FT      /tag= h
FT      /product= "spnN"
FT      /note= "Spinosyn biosynthesis protein N"
FT      complement(11530..12492)
FT      /tag= i
FT      /product= "spnM"
FT      /note= "Spinosyn biosynthesis protein M"
FT      complement(12656..13547)
FT      /tag= j
FT      /product= "spnL"
FT      /note= "Spinosyn biosynthesis protein L"
FT      complement(13592..14785)
FT      /tag= k
FT      /product= "spnK"
FT      /note= "Spinosyn biosynthesis protein K"
FT      complement(14799..16418)
FT      /tag= l
FT      /product= "spnJ"
FT      /note= "Spinosyn biosynthesis protein J"
FT      16556..17743
FT      /tag= m
FT      /product= "spnI"
FT      /note= "Spinosyn biosynthesis protein I"
FT      complement(17749..18501)
FT      /tag= n
FT      /product= "spnH"
FT      /note= "Spinosyn biosynthesis protein H"
FT      complement(18541..19713)
FT      /tag= o
FT      /product= "spnG"
FT      /note= "Spinosyn biosynthesis protein G"
FT      20168..20995
FT      /tag= p
FT      /product= "spnF"
FT      /note= "Spinosyn biosynthesis protein F"
FT      21111..28898
FT      /tag= q
FT      /product= "spnA"
FT      /note= "Spinosyn biosynthesis protein A"
FT      /transl_except= (Pos:26940..26942, aa:Pro)
FT      21126..22379
FT      /tag= r
FT      /note= "Beta-ketosynthase domain (KS1): part of the
FT      initiator module"
FT      22692..22669
FT      /tag= s
FT      /note= "Acyl transferase domain (AT1): part of the
FT      initiator module"
FT      23793..24041
FT      /tag= t
FT      /note= "Acyl carrier protein domain (ACPl): part of the
FT      initiator module"
FT      24102..25349
FT      /tag= u
FT      /note= "Beta-ketosynthase domain (KS1): part of extender
FT      module 1"
FT      25683..26684
FT      /tag= v
FT      /note= "Acyl transferase domain (AT1): part of extender
FT      module 1"
FT      27582..28121
FT      /tag= w
FT      /note= "Beta-ketoreductase domain (KR1): part of extender
FT      module 1"
FT      28404..28649
FT      /tag= x
FT      /note= "Acyl carrier protein domain (ACPl): part of
FT      extender module 1"
FT      29024..30295
FT      /tag= z
FT      /note= "Beta-ketosynthase domain (KS2): part of extender
FT
```


FT		module 2"
FT	misc_feature	30629. .31621
FT		/*tag= aa
FT		/note= "Acyl transferase domain (AT2) : part of extender
FT		module 2"
FT	misc_feature	31697. .32524
FT		/*tag= ab
FT		/note= "dehydratase domain (DH2) : part of extender module
FT		2"
FT	misc_feature	33035. .34072
FT		/*tag= ac
FT		/note= "hnoyl reductase domain (ER2) : part of extender
FT		module 2"
FT	misc_feature	34082. .34621
FT		/*tag= ad
FT		/note= "beta-ketoreductase domain (KR2) : part of extender
FT		module 2"
FT	misc_feature	34886. .30295
FT		/*tag= ae
FT		/note= "Acyl carrier protein domain (ACP2) : part of
FT		extender module 2"
FT	CDS	35419. .44931
FT		/*tag= af
FT		/product= "spnC"
FT		/note= "Spinovyn biosynthesis protein C"
FT	misc_feature	35518. .36786
FT		/*tag= ag
FT		/note= "beta-ketosynthase domain (KS3) : part of extender
FT		module 3"
FT	misc_feature	37108. .38097
FT		/*tag= ah
FT		/note= "Acyl transferase domain (AT3) : part of extender
FT		module 3"
FT	CDS	38916. .35374
FT		/*tag= y
FT		/product= "spnC"
FT		/note= "Spinovyn biosynthesis protein C"
FT	misc_feature	38992. .39528
FT		/*tag= ai
FT		/note= "beta-ketoreductase domain (KR3) : part of extender
FT		module 3"
FT	misc_feature	39790. .40035
FT		/*tag= aj
FT		/note= "Acyl carrier protein domain (ACP3) : part of
FT		extender module 3"
FT	misc_feature	40102. .41373
FT		/*tag= ak
FT		/note= "beta-ketosynthase domain (KS4) : part of extender
FT		module 4"
FT	misc_feature	41713. .42705
FT		/*tag= al
FT		/note= "Acyl transferase domain (AT4) : part of extender
FT		module 4"
FT	misc_feature	43615. .44157
FT		/*tag= am
FT		/note= "Beta-ketoreductase domain (KR4) : part of extender
FT		module 4"
FT	misc_feature	44431. .44676
FT		/*tag= an
FT		/note= "Acyl carrier protein domain (ACP4) : part of
FT		extender module 4"
FT	CDS	44966. .59752
FT		/*tag= ao
FT		/product= "spnD"
FT		/note= "Spinovyn biosynthesis protein D"
FT	misc_feature	45077. .46348
FT		/*tag= ap
FT		/note= "Beta-ketosynthase domain (KS5) : part of extender
FT		module 5"
FT	misc_feature	46691. .47674
FT		/*tag= aq
FT		/note= "Acyl transferase domain (AT5) : part of extender
FT		module 5"

FT	misc_feature	47753..48310	
FT		/+tag= ar	
FT		/note= "dehydratase domain (DH5) : part of extender module 5"	
FT	misc_feature	49226..49771	
FT		/+tag= ag	
FT		/note= "beta-ketoreductase domain (KR5) : part of extender module 5"	
FT	misc_feature	50009..50254	
Alignment Scores:			
Pred. No.:	6.24e-31	Length:	80161
Score:	435.50	Matches:	108
Percent Similarity:	51.67%	Conservative:	47
Best Local Similarity:	36.00%	Mismatches:	96
Query Match:	30.16%	Indels:	49
DB:	2	Gaps:	11
US-10-069-353A-8 (1-275) x AA221501 (1-80161)			
QY	7	ProthrSerGlnGlnValGlyGlnMetYrAspLeuValThrProLeuLeuAnsSerVal	26
DB	13562	CCGACCGCGGATCAGGTGAGGTCGATCTTCGATGCGTTG-----GCG	135211
QY	27	AlaGlyLyrProCysAlaIleHisGlyYrTrp-----GluLeuAsp	41
DB	13520	CACGGGCGTCCC-----CTGCACACAGGTTACTGGGCGGCGGCGGATGCGGAGATGCC	134675
QY	42	GlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArgThr	61
DB	13466	GGTGCCACACCGTGTGCGATGCTGCCGACCACTACGACCTGTTCATCGACAAGGCC	134070
QY	62	ValLeuAspArgGlyValAlaArgLeuLeuAspValGlyCysGlyThrGlyGlnProAlaLeu	81
DB	13406	CGCGTCCGTCGCCGAGCGCACCTGTTGCACCTGGGCTGGCGCAATGGGACGCCGATGTC	133474
QY	82	ArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGlnVal	101
DB	13346	CGTCCGCGCATGCGCCAGCGCGCTTCGATGTCACCGGATATCACCGTGAACGCCAGCATCTC	132875
QY	102	AlaIleIleAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSerCys	121
DB	13286	GCCGCGCGCACCGAGCTGCGCAACGAGACCGGACTGCGGCACTTGAHTTCATCTA	132222
QY	122	ValAspAlaMetSerLeuProYrProAspAsnAlaPheAspAlaAlaTrpAlaMetGln	141
DB	13226	GTCACGCGGCCCGCCAGCTGCCCTTACCGGACGGTTCCTTCAGCGCCGACATGGGAGTGCAG	131676
QY	142	SerLeuLeuGlnMetSerGlnProAspArgAlaIleArgGluIleLeuArgValLeuLys	161
DB	13166	TCCGTCGTCGAGATCTGTCAGCACAGCGCGCGCATCCGCGACGTCACCGAATCCCTGGAA	131070
QY	162	ProGlyGly-----IleLeuGly-----ValThrGluValValLysArgGluAlaGly	177
DB	13106	CCCGCGCGCGGTTGCTCTCGGAGACATCATCATCTGGGTTGACTCCCGGAAGCATTC	130474
QY	178	GlyLeuMetProValSerGlyAspArgYrTrpProThrGlyLeuArgIleCysLeuAlaGlu	197
DB	13046	CGCGCGGTT-----TGG-----ACGGGACACGACCGCCCATACCTTGAAC	130086
QY	198	GlnLeuLeuGlnSerLeuArgAlaAlaGlyPheGlnIleLeuAspTrpGluAspValSer	217
DB	13007	AGCTTCACCGGCGTGTGTCAGCGAAGCGGGGTTGAGATTCGAAATGTCACCGACCTTCAGG	129448
QY	218	SerArgThrArgYrYrPheMetProGlnPheAlaGluGlnLeu-----	231
DB	12947	GCAACAGACAGTGCATGTCTCTCTGTATGTCGACAGATGTTCTCCGAAACTCGATGAG	128866
QY	232	-----AlaAlaHisGlnHis-----GlyIle	238
DB	12887	CTCCGCGCGCGTCGACGCTCGCGCTGCGGACCTTACCGAAGACGATCTTGGGAGACATC	128288
QY	239	AlaAspArgYrGlyProAlaValAlaGlyTTrpAlaAlaAlaValCysAspYrGlyLys	258


```
Db      12827 GCGGCGAAGCAGGACCGGACGACGATGATCGCGCGGTTCGGGAATACCGGAAA 12768
      259 -----TyrAlaHisasp-----MetGlyTyrAlaIleLeuThrAlaArglys 272
Db      12767 CATCGGATATACCGCAGAAACGAGAAAGCATGGTTTCCTCTCGAGGCTCGAAG 12708

RESULT 14
AAT70153
ID      AAT70153 standard; DNA; 6085 BP.
XX
XX      AAT70153;
XX
XX      17-OCT-2003 (revised)
DT      23-JUL-1997 (first entry)
XX
XX      S.longisporoflavus staurosporin synthesis gene cluster 6.5kb fragment.
DE
XX      Staurosporin; gene cluster; indole-carbazole alkaloid; ICA; antibiotic;
KW      antiproliferative; platelet aggregation; fungus; yeast;
KW      Ca2+/phospholipid-dependent serine/threonine protein kinase; ds.
XX
OS      Streptomyces longisporoflavus; (strain R19).
XX
XX      Key      Location/Qualifiers
XX      CDS      378..1655
XX              /tag= a
XX              /label= Gene 1
XX              /note= "Encodes a protein containing 425 amino acids"
XX
XX      CDS      1747..2553
XX              /tag= b
XX              /label= Gene 2
XX              /note= "Encodes a protein containing 268 amino acids
XX              significantly similar to known S-adenosyl methionine-
XX              dependent methyl transferase"
XX              2593..4011
XX              /tag= c
XX              /label= Gene 3
XX              /note= "Encodes a protein containing 472 amino acids"
XX              4013..4999
XX              /tag= d
XX              /label= Gene 4
XX              /note= "Encodes a protein containing 328 amino acids"
XX              5071..6085
XX              /tag= e
XX              /label= Gene 5 (part)
XX              /note= "Partial sequence for gene 5; full length protein
XX              contains 366 amino acids and is significantly similar to
XX              amino transferase enzymes, such as the Dnr J protein"
XX
XX      CDS
XX              /tag= e
XX              /label= Gene 5 (part)
XX              /note= "Partial sequence for gene 5; full length protein
XX              contains 366 amino acids and is significantly similar to
XX              amino transferase enzymes, such as the Dnr J protein"
XX
XX      WO9708323-A1.
XX
XX      PD      06-MAR-1997.
XX
XX      PF      19-AUG-1996; 96WO-EP0036643.
XX
XX      PR      30-AUG-1995; 95EP-00810534.
XX
XX      PA      (CIBA) GEIGY AG.
XX
XX      PI      Schupp T, Engel N, Bietenhader J, Toupet C, Pospiesch A;
XX
XX      WPI; 1997-179280/16.
XX
XX      DR      Indole-carbazole alkaloid biosynthesis gene cluster - especially coding
XX      PT      for the antibiotic staurosporin from Streptomyces longisporoflavus.
XX
XX      PS      Claim 7; Page 37-41; 55pp; English.
XX
XX      CC      The present sequence represents the 6.5kb PvuII fragment of Streptomyces
XX      CC      longisporoflavus R19 which is involved in the biosynthesis of indole-
XX      CC      carbazole alkaloids (ICA). The sequence contains five functional
XX      CC      fragments as indicated in the features table. The DNA or a hybrid vector
```

```
CC      containing it can be used to prepare an ICA or derivatives and
CC      precursors, either by allowing production in previously incapable
CC      organisms or by improving yields. In particular, the antibiotic
CC      staurosporin can be produced. Staurosporin is known to have inhibitory
CC      activity against fungi, yeasts, and Ca2+/phospholipid-dependent
CC      serine/threonine protein kinases (PKCs). Staurosporin also has
CC      antiproliferative activity and can inhibit platelet aggregation. The
CC      present sequence can also inactivate ICA biosynthesis genes and can be
CC      used in PCR amplification. An advantage of this is that productivity of
CC      staurosporin-synthesising Streptomyces is improved over natural strains
CC      yielding only low concentrations. (Updated on 17-OCT-2003 to standardise
CC      OS field)
```

Sequence 6085 BP; 882 A; 2374 C; 1992 G; 833 T; 0 U; 4 Other;

Alignment Scores:

```
Pred. No.: 9.17e-27 Length: 6085
Score: 379.00 Matches: 101
Percent Similarity: 51.44% Conservative: 42
Best Local Similarity: 36.33% Mismatches: 113
Query Match: 26.25% Indels: 23
DB: 2 Gaps: 4
```

US-10-069-353A-8 (1-275) x AAT70153 (1-6085)

```
Qy      8 ThrSerGlnGlnValGlyGlnMetTyrAspLeuValThrProLeuLeuAsnSerValAla 27
      1768 ACCGCCAAGTCGTCGCGAGAGTGTACACAGTTCGCCAGCGCGCGCCAGACCGCG 1827
Qy      28 GlyGlyProCysValAlaIleHisGlyTyrTrpGlnAsnArgIlyArgAlaSerTrpGln 47
      1828 ATGGCGCGC--AACATCCAGTGGGGTACTGGAGACGACCCCGAGTCCGATCGCC 1884
Qy      48 GlnAlaAlaAspArgValLeuThrAspLeuValAlaGlnArgTrpValIleuAspGlyVal 67
      1885 GAGGCCACGACCGAGCTACCGATCTGTCGCGCAGGCGCTCGCGCTCCGCCGACCGG 1944
Qy      68 ArgLeuLeuAspValGlyCysGlyThrGlyGlnProAlaLeuArgValAlaArgAspAsn 87
      1945 CATCTGCTGAGACGTCGCGTGGCGATCGCGCTCGCGCTCTCAGAGATCGCGGACCGAC 2004
Qy      88 AlaIleGlnIleThrGlyIleThrValSerGlnValGlnValAlaIleAlaAspCys 107
      2005 GACGTCGCGTCACCGGATACCGTCAGCGACGAGGATCACCGGCGCGCGCGCGCG 2064
Qy      108 AlaArgGlnArgGlyLeuSerHisArgValAspHisSerCysValAspAlaMetSerLeu 127
      2065 GCGGTGAGTCCGATGCCGCGCGCGCGCGCTCTCCGCGTGGCGGACCGCATGACCTC 2124
Qy      128 ProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMetGlnSerLeuLeuGlnMetSer 147
      2125 CCTTCGAGGACGTCCTCTTGAAGCGCGCTTCGCGCATTCAGTCTGCTCATCTGCC 2184
Qy      148 GluProAspArgAlaIleArgGlnIleLeuArgValLeuLysProGlyGlyIleLeuGly 167
      2185 GACGAGACACCGCGCTCAGAGAGATCACCGGCTGTCGCGCGCGCGCGCGCTGCTC 2244
Qy      168 ValThrGlnValValValArgGlnAlaGlyGly----- 179
      2245 ATCGCGACCTGTGTACGACGACCGCTTCAACCGCGCGCGCGCGCGAGAGTCTGACGG 2304
Qy      180 -----MetProValSerGlyAspArgTrpProThrGlyLeuArgIleCysLeu 195
      2305 ATGCTGCTGATGATGAGAGATCGCGGATCAACACACCTTACGAGATCGCGCGCGCATG 2364
Qy      196 AlaGlnLeuLeuLeuGlnSerLeuArgAlaAlaGlyPheGlnIleLeuAspTrpGlnAsp 215
      2365 GCGGAG-----GCGGCGTGGGAACTGCTGAGCTGACCGGAC 2400
Qy      216 ValSerSerArgThrArgTrpPheMetProGlnPheAlaGlnGlnLeuAlaIleGln 235
      2401 ATCGGTAGACGAGTTCGCGCTACTTACGCGGACATCCGCCCGC-CGCGTTCGCGGCTCTGC 2459
```

QY 226 HIGIYLLAALAPATGTYGlyProAlaValAlaGlyTPAlaAlaValCysAsp 255
DB 2460 CGGGGCTCTGACGCGCGCGCGACGAGTGAACGCGCGGCGGCTGATGAGGC 2519
QY 256 TYRGlyLysTYRAlaHisAspMetGlyTYRAlaIleLeuThrAlaArgLysPro 273
DB 2520 TTCGAGGGGATCCGACATCC--GGTTCGTCCTGATGACGCGGCGGTCT 2570
RESULT 15
AAC55802
ID AAC55802 standard; DNA; 852 BP.
AC AAC55802;
AT 19-JAN-2001 (first entry)
DE S. Javendulae Mitc encoding DNA sequence.
KW Mitomycin; biosynthesis; mitosome ring system; antibiotic; anti-cancer;
KW anti-inflammatory; immune-enhancer; immunosuppressant; asthma;
KW chronic obstructive pulmonary disease; respiratory inflammation;
KW fungicide; pesticide; ds.
OS Streptomyces Javendulae.
PN MO200053737-A2.
XX 14-SEP-2000.
XX 10-MAR-2000; 2000MO-US006394.
XX 12-MAR-1999; 99US-00266965.
PA (MINU) UNIV MINNESOTA.
PA (SHEM) SHERMAN D H.
PA (MAOY) MAO Y.
PA (VARO) VAROGU M.
PA (HEMM) HE M.
PA (SHEL) SHELTON P C.
PI Sherman DH, Mao Y, Varoglu M, He M, Sheldon PC;
DR WPI, 2000-601980/57.
DR P-PSDB; AAB32457.
XX
PT Novel nucleic acid molecule comprising mitomycin biosynthetic gene
PT cluster useful for cloning mitomycin biosynthetic genes for elucidating
PT the molecular basis of mitosome ring system biosynthesis.
XX
XX Example 1; Page 262-263; 399pp; English.
XX
CC This invention relates to isolated and purified nucleic acid molecules
CC from the mitomycin biosynthetic gene cluster. Mitomycins are a group of
CC natural products that contain a variety of functional groups, including
CC amino benzoxiquone and axiridine ring systems. The S. Javendulae
CC mitomycin biosynthetic gene cluster comprises 47 mitomycin genes spanning
CC 55kb of DNA. The invention includes an expression cassette comprising a
CC mitomycin biosynthetic gene operably linked to a promoter, and host cells
CC transformed with the cassette. The nucleotide, and protein sequences and
CC the transformed host cells of the invention result in antiasthmatic,
CC anti-inflammatory, cytostatic, immunomodulatory, and antibiotic
CC activities. The nucleotide sequences are used to elucidate the molecular
CC basis for the biosynthesis of the mitosome ring system, as well as to
CC engineer the biosynthesis of novel natural products, e.g. antibiotics,
CC anti-inflammatory agents, anti-cancer agents, immune-enhancers,
CC immunosuppressants, agents to treat asthma, chronic obstructive pulmonary
CC disease as well as other disease involving respiratory inflammation, or
CC cholesterol-lowering agents or as crop protection agents (e.g. fungicides
CC or insecticides) as well as biopolymers, e.g., in packaging or biomedical
CC applications, or to engineer PHA monomer synthases. Sequences AAC55782-
CC C55881, AAC55815-C55849 and AAB32485-832542 represent mitomycin
CC biosynthetic gene cluster DNA sequences and encoded proteins. Sequences
CC AAC55812-C55814, AAC55850-C55856 and AAC55862-C55869 represent PCR

CC primers used in the cloning of the mitomycin biosynthetic genes
XX
SQ Sequence 852 BP, 125 A, 316 C, 266 G, 125 T, 0 U, 0 Other;
Alignment Scores:
Pred. No.: 1,386-27 Length: 852
Score: 376.50 Matches: 101
Percent Similarity: 50.84% Conservative: 51
Best Local Similarity: 33.78% Mismatches: 94
Query Match: 26.07% Indels: 53
DB: Gaps: 11
US-10-069-353A-8 (1-275) x AAC55802 (1-852)
QY 2 LeuProGlyGlyAlaProThrSerGlnIleValGlyGlnMetTYRAspLeuValThrPro 21
DB 25 CTCGCCATGCCCTCACCACCGCGCTCCGAGAGTGGCGCGCTTACGACCGGTCCACGCG 84
QY 22 LeuLeuAsnSerValAlaGlyGlyProCysAlaIleHisAspGlyTYRTrpGluAsn-- 40
DB 85 CTGGAGCCGCGCTCCCTCGGCGAG-----ACCTGCACTTCGGCTACGGAGATCCCCC 138
QY 41 AspGlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
DB 139 GACAGCGAGTGCCTCGTGGCGAGGCCACCGGCTCACCGACATGATGCGCGAGCGCG 198
QY 61 ThrValLeuAspGlyGlyValAlaArgLeuLeuAspValGlyCysGlyTYRTrpGlyProAla 80
DB 199 CTGCGCATCGCGCGCGCTCCGCGCTTCGACCTCGCTGCGCGGTGGAGACCCCGGCG 258
QY 81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyTleThrValSerGlnValGln 100
DB 259 GTACGCATGCCCGGCTCCACGAGCGCATGTGCACGGGATCTCGGTGAGCATGAGCGAG 318
QY 101 ValAlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120
DB 319 GTCTGTCGGGCGCAACGCGGTGGCCGAGAGCGCGGCTCGCGACCGCGCGCGCTTCGAG 378
QY 121 CysValAspAlaMetSerLeuProTYRProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
DB 379 CGGCGCGAGCGCATGACCTCCCTTCGAGAGCGAGAGCTTGACCCCTCATCGCCCTTC 438
QY 141 GlnSerLeuGlnLysSerGluProAspArgAla-----IleArgGlnIleLeuArg 158
DB 439 GAATGCATCATCCACATG-----CCGACCGCGCCAGGTGTCGCCACGATCGGCCG 492
QY 159 ValLeuLysProGlyGlyTleLeuGlyValThrGluValValArg----- 174
DB 493 GTCTCGGCGCGCGGCGCGCTGCTGTCGACCGGACTTCTTCGAGCGGCGCCCTCGCC 552
QY 175 --GluAlaGlyGlyGlyMetPro-----ValSer 183
DB 553 CCCGAGGGGCGCGCGCGCTCCAGCGCTTACGACTTCATGATGACCATGATGATGACG 612
QY 184 GlyAspArgTrpProThrGlyLeuArg-----IleCys 194
DB 613 GCCGAGCGGATCCCTCCCTCGCTCGCGGCGCGGCGCTGCTGCTGAGAGAGTTCCTGCAC 672
QY 195 LeuAlaGlnIleLeuGlnLysSerLeuArgAlaAlaGlyPheGlnIleLeuAspTrpGlu 214
DB 673 ATGAGGAGCAGACACCTCGGAGAGAC-----TTGAGCGTGGCTCTCG--GAG 717
QY 215 AspValSerSerArgThrArgTYRPhenMetProGlnPheAlaGlnIleLeuAlaHis 234
DB 718 CGCATCAACTCTTCGAGAGCGAGCTGAGACCGCATTTGCGCGAGAGATG--GTGAC 774
QY 235 GlnHisGlyTleAlaAspArgTYRglyProAlaValAlaGlyTPAlaAlaValCys 254
DB 775 CAGTTGACCCCGCGGACCTCGTGGC----- 801
QY 255 AspTYRgluLysTYRAlaHisAspMetGlyTYRAlaIleLeuThrAlaArgLysPro 273
DB 802 -----GTCAAGAGATTGCGGTATCTGCTGTGTCGCCCGAGCCCGG 843

Search completed: August 29, 2005, 20:03:36
Job time : 747 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 29, 2005, 19:36:04 ; Search time 3539 Seconds
(without alignments)
2957.806 Million cell updates/sec

Title: US-10-069-353A-8

Perfect score: 1444
Sequence: 1 VLPGAPTSQVGVQWVLT.....YKRYAHMGVAILTPKPVG 275

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 6847908

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-O/cgnt2.1/USPTO.spool/US1006953/rnatc.26082005.172054.14678/app.query.fasta_1.455
-DB=EST -QPM=fastap -SUFFIX=ret -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=numan40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-USER=US1006953 @CGN 1.1 3437 @rnatc.26082005.172054.14678 -NCPU=6 -ICPU=3
-NO_MMP -LARGEJURY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	263.5	18.2	787	6	CA764573 AP53-RP6
2	245.5	17.0	842	7	CN814661 HRO4501 B
3	241	16.7	575	5	B0514932 EST622347
4	241	16.7	859	7	CO076074 GR_EA37D
5	240.5	16.7	980	7	CO008962 WHE2647 B
6	240.5	16.7	980	7	CO009336 EST797671
7	240	16.6	824	7	CO129317 GR_Eb260
8	239.5	16.6	902	7	CF515375 CAP0001 I
9	238.5	16.5	669	6	CA220497 SCRUF1402

10	238.5	16.5	1665	3	AY103656
11	237.5	16.4	684	7	CA927208
12	236.5	16.4	786	6	CO120890
13	236.5	16.4	961	7	CO009335
14	236	16.3	634	1	AJ431900
15	234	16.2	650	4	BJ271067
16	231.5	16.0	1001	7	CO004512
17	231	16.0	643	5	B0408882
18	230.5	16.0	723	7	CF476020
19	230.5	16.0	755	6	CA064977
20	230.5	16.0	811	4	BM405193
21	230.5	16.0	815	4	BM405276
22	230	15.9	814	7	CN812339
23	229	15.9	854	7	CO092359
24	229	15.9	624	2	AM424526
25	228.5	15.8	592	5	B0990099
26	228	15.8	642	6	CA284871
27	227.5	15.8	592	6	CA101641
28	227.5	15.8	628	2	BF278472
29	227	15.7	600	5	BU640731
30	226.5	15.7	626	6	CD037447
31	226.5	15.7	842	7	CO086997
32	226.5	15.7	916	7	CO082474
33	226	15.7	861	7	CO117413
34	226	15.7	891	7	CN809395
35	225.5	15.6	688	2	BF479484
36	225	15.6	849	7	CO070289
37	224	15.5	858	7	CK935285
38	222.5	15.4	952	7	CO082835
39	221.5	15.3	1198	3	CN50A23V
40	221	15.3	751	6	CB289364
41	219.5	15.2	629	6	CA170495
42	218.5	15.1	771	7	CO367476
43	218	15.1	547	2	BF598669
44	218	15.1	587	4	BG363118
45	217	15.0	570	2	AW706079

ALIGNMENTS

RESULT 1
CA764573
LOCUS
DEFINITION
CA764573
AP53-RP6 04 B21.T7 081.ab1 IIRI Drought Stress Panicle cDNA library
Oryza sativa (indica cultivar-group) cDNA clone C0001197.5, similar to Delta (24)-sterol C-methyltransferase (EC 2.1.1.41), mRNA sequence.

ACCESSION
CA764573
VERSION
CA764573.2 GI:27546528
KEYWORDS
EST.
SOURCE
Oryza sativa (indica cultivar-group)
ORGANISM
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
AUTHORS
Bennett,J., Arumugam,K., Lafitte,R., Wen,J., Rudd,S. and Bruskiewicz,R.M.
TITLE
IIRI Drought Stress Panicle cDNA library
JOURNAL
Unpublished (2002)
COMMENT
On Dec 2, 2002 this sequence version replaced gi:25993828.
Contact: Richard Bruskiewicz
Biometrics and Bioinformatics Unit
International Rice Research Institute
DAPO 7777, Metro Manila, Philippines
Tel: +63-2-845-0563
Fax: +63-2-845-0606
Email: r.bruskiewicz@icrri.org
International Rice Information System (IRIS)
http://www.iris.icrri.org: D0201196
Assessment of putative function to the sequence by S. Rudd of the Munich Information Center for Protein Sequences
(http://mips.gsf.de)


```
Db      128 GAATCTTTTCACTTGTCTAC---AGTGGAAATGGGAA-----TCTTTAAGGAA 175
Qy      49 AlaAlaapAargLeuthrAapLeuValAlaGluArgThrValLeuAapGlyValArg 68
Db      176 AGATCAAAACGACAGAGCATTTTCTGGCCCTTCAGCTTGAGTGAACAGGAGTAGAA 235
Qy      69 LeuLeuAapValGlyCysGlyThrGlyGlnProAlaLeuArgValAlaArgAapAla 88
Db      236 GTTTTGACCTGGCGCTGTGGAATAGGTGACCTTAAGAGATTCACGATTACTCTCA 295
Qy      89 IleGlnIlethrgGlyIleThrValSerGlnValGlnValAlaAlaIleAlaAapCysAla 108
Db      296 ACCCTGATTACCGGATTGGAACAACAGCACTATCAGTACTGAGGGAAGAGCTCAAT 355
Qy      109 ArgGluArgGlyLeuSerHisArgValAspPheSerCysValAlaMetSerLeuPro 128
Db      356 CGTCTGGCAGGACTTGTAGTGAATCTTGTCAAGCAGACTTCATGAAATATGCTT 415
Qy      129 TyrProAapAapAlaPheAapAlaAlaTPrAlaMetGlnSerLeuGluMetSerGlu 148
Db      416 TTCTCCGATTAACACTTTTGTATGCTGTATATGATGAGCAACATGTCATGACCTGAT 475
Qy      149 ProAapAargAlaIleAargGluIleAargValIleuArgProGlyGlyIleuGlyVal 168
Db      476 CCGGTTGGCTGTATTAAGAGATCTACCGGTATTAACCTGGGACGCTTTGCTGTA 535
Qy      169 ThrGlu-----ValVallyArg----- 174
Db      536 TATGAGTGGCTACTTACTGATCACTATGATCCAAACATGCAACCAAGAGATTAAG 595
Qy      175 -----GluAlaGlyGlyGlyMetProValSerGlyAapArgTrpProthrgGlyLeu 191
Db      596 GATGAATATGAGCTGTGTAATGCTTACCA----- 625
Qy      192 ArgIleCysLeuAlaGluGlnLeuGlnSerLeuAargAlaIleGlyPheGluIleu 211
Db      626 GATATCAAGAACTCAAGCAATGCTTCAAGCTGTTAAGATGAGGATTTAGAGTATAT 685
Qy      212 AspTrpGluAap-----ValSerSerAargThrArgTrpPheMetPro 225
Db      686 ----TGGGATTAAGATCTTGGCTGTATGATTCATTCGCTTGGTACTTGCC 733

RESULT 3
BO514932      575 bp      mRNA      linear      EST 07-MAR-2003
LOCUS      BO514932
DEFINITION      EST622347 Generation of a set of potato cDNA clones for microarray
                    analyses mixed potato tissues Solanum tuberosum cDNA clone STMIP11
                    5' end, mRNA sequence.
ACCESSION      BO514932
VERSION      BO514932.1 GI:21373801
KEYWORDS      EST.
SOURCE      Solanum tuberosum (potato)
ORGANISM      Solanum tuberosum
REFERENCE      Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C.,
ATTNORS      Reestrop, S., Griffiths, H., van der Hoeven, R., Tesi, J. and
                    Karayancheva, S.A.
                    Generation of a set of potato cDNA clones for microarray analyses
                    Unpublished (2002)
                    Other ESTs: EST622348
                    Contact: Robin Buell
                    The Institute for Genomic Research
                    9712 Medical Center Dr., Rockville, MD 20850, USA
                    Email: potato-array@tigr.org
                    This clone can be obtained from the University of Arizona Genomics
                    Institute. Orders can be made through URL:
                    http://genome.arizona.edu/orders/
                    Seq primer: T3.
FEATURES      Location/Qualifiers
```

```
..      source      1..575
                    /organism="Solanum tuberosum"
                    /mol_type="mRNA"
                    /cultiivar="Kennebec or Binje"
                    /db_xref="taxon:4113"
                    /clone="STMIP11"
                    /tissue_type="mixed tissues"
                    /lab_host="SOLR"
                    /clone_lib="Generation of a set of potato cDNA clones for
                    microarray analyses mixed potato tissues"
                    /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
                    XhoI; supplier: Combination of untreated and Phytophthora
                    infestans-created libraries of stolons, leaves, leaflets,
                    axillary buds of stem explants, petioles, germinating
                    eyes, tubers, or roots."

ALIGNMENT SCORES:
Pred. No.:      9,52e-16      Length:      575
Score:      241.00      Matches:      64
Percent Similarity:      48.73%      Conservative:      32
Best Local Similarity:      32.49%      Mismatches:      29
Query Match:      16.69%      Indels:      29
DB:      5      Gaps:      4

US-10-069-353A-8 (1-275) X BO514932 (1-575)

Qy      45 SerTrpGlnGlnAlaAlaAapAargLeuthrAapLeuValAlaGluArgThrValLeuAap 64
Db      7 TCACCTCAGGAAGACTTAAGAGCATGACGACTTCCTTCCTTCGCAATTCGGAATGAA 66
Qy      65 GlyGlyValArgLeuLeuAapValGlyCysGlyThrGlyGlnProAlaLeuArgValAla 84
Db      67 CCAGCAAAAGCTCTGTGATGTAGAGATGGAATTGAGGCCCTTAAGAAATGCT 126
Qy      85 ArgAapAapAlaIleGlnIlethrgGlyIleThrValSerGlnValGlnValAlaIleAla 104
Db      127 CATTCAAGCTCTACATCACTTACAGGCTTCACACAAATGAAATACAGTATCTAGCGGA 186
Qy      105 AlaAapCysAlaArgGluArgGlyLeuSerHisArgValAspPheSerCysValAapAla 124
Db      187 CAGGTGTGAACCGCAAGTGGGATTTGATCAAACTTGCACTTTGTAAGGGTGTATTC 246
Qy      125 MetSerLeuProTyrProAapAapAlaPheAapAlaAlaTPrAlaMetGlnSerLeu 144
Db      247 ATGAATAATCCATTCTCTGCAATAGCTTGTATGCACTATGACATAGAGCTACTGTC 306
Qy      145 GluMetSerGluProAapAargAlaIleAargGluIleLeuArgValIleuArgProGlyGly 164
Db      307 CATGACCCAGATCCGTTGGATGCTATAGAGAGATCTATAGGCTGTGAAGCTGTGTCAG 366
Qy      165 IleLeuGlyVal-----ThrGluValVal 172
Db      367 TGCTTGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 426
Qy      173 LysArg-----GluAlaGlyGlyGlyGlyMetProValSerGlyAapArgTrp 187
Db      427 AAAAGATCAAGAAAGAAATTAAGCTCGGAAATGCTTCG----- 468
Qy      188 ProthrgGlyLeuArgGlyCysLeuAlaGluGlnLeuGlnSerLeuAargAlaIleGly 207
Db      469 -----GAGATTCGATCGACACAAAGTGTCTTGAAGCAGCGGAAGAGCTGCT 516
Qy      208 PheGluIleLeuAapTrpGluAap-ValSerSerAargThrArgTrpPhe 223
Db      517 TTTGAAGTTGTG---TGGGATTAAGATCTCGCTGAAGACTCACCTGTTT 562

RESULT 4
CO076074      859 bp      mRNA      linear      EST 15-JUN-2004
LOCUS      GR_Ea37D20.r GR__Ea Gossypium raimondii cDNA clone GR__Ea37D20 3',
DEFINITION      mRNA sequence.
ACCESSION      CO076074
```

```

VERSION      CO076074.1  GI:48745555
KEYWORDS     EST.
SOURCE       Gosyplium raimondii
ORGANISM     Gosyplium raimondii
REFERENCE    1 (bases 1 to 859)
AUTHORS      Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
              Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
              Ming,R.A.
TITLE        Global assembly of Cotton ESTs
JOURNAL      Unpublished (2004)
COMMENT      Arizona Genomics Institute
              The University of Arizona
              Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
              Tel: 520 626 9595
              Fax: 520 621 1259
              Email: http://genome.arizona.edu
              Plate: 37 row: D column: 20.
              Location/Qualifiers
FEATURES
    source          1..859
                    /organism="Gosyplium raimondii"
                    /mol_type="mRNA"
                    /db_xref="taxon:29730"
                    /clone="GR_Ea37D20"
                    /tissue_type="whole seedlings"
                    /dev_stage="first true leaves"
                    /lab_host="DHIOB"
                    /clone_lib="GR_Ea"
                    /note="Vector: PCWV.SPORT-6.1; Site 1: NotI; Site 2:
                    EcoRV; Library made by Invitrogen with RNA supplied by
                    Wendle lab. Directional cloned into NotI-EV. Colonies
                    plated/picked by AGI. More glycerol clones held in -80."

ORIGIN
Alignment Scores:
Pred. No.:      1..65e-15      Length:      859
Score:          241.00         Matches:     69
Percent Similarity: 43.15%      Conservative: 38
Best Local Similarity: 27.82%      Mismatches:  107
Query Match:    16.69%         Indels:       34
                    Gaps:        8

US-10-069-353A-8 (1-275) x CO076074 (1-859)
QY      3  ProGlyGlyAlaProThrsSerGlnGlnValGlyGlnMetTyrAspLeuValThrProLeu 22
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      103 CCCGATGCACGTGATTCGACGAGGAGGAGCGAGGTTCTACGACCGATCGTCGGGCTT 162
QY      23  LeuAanSerValAlaGlyGlyProCysAlaIleHisIleGlyTyrTyrProGluAanSpGly 42
        ::|  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      163 TGGGAAGACATTTGGGGTGAC-----CATATGATCATGATTTTACCATCGGATTTCC 216
QY      43  ArgAlaSerTyrGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArgThrVal 62
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      217 AATGTTTCAGGCTCCGATCATCCAGCTGCCAGATCCGAATGATC--GAAGAATCGGTC 273
QY      63  LeuAanSpGlyVal-----ArgLeuLeuAanSpGlyValGly 73
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      274 CGTTTCGCTGAATACCGATATCCAGCAAGACGCCCAAGACAAATATGATTTGGA 333
QY      74  CysGlyThrGlyGlnProIleLeuArgValAlaArgAspAlaAlaGlnIleThrGly 93
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      334 TGTGGATAGAGGACGCTCTAGATACCTTAGCAAGGAAATTTGGAGCAGAAATGCTAAGGC 393
QY      94  IleThrValSerGlnValGlnValAlaIleAlaAspCysAlaArgGluArgGlyLeu 113
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      394 ATTACTTTCAGCCCTGTCCCAAGCTGGAAGCGCTAATGTTCTTGCTTAAGATGAAGACGTA 453
QY      114 SerHisArgValAlaSpHesSerCysValAspAlaMetSerLeuProTyrProAspAla 133
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

```

Db      454 GCAGCAAGGTTTCATTTCAAGTTGCAGATGCTCTGAAGCAACCATTTCCGTAGTGCAG 513
QY      134  PheAspAlaIleTPrAlaMetGlnSerLeuGluMetSerGlnProAspArgAlaIle 153
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      514 TTGATCTAGATTGGCTATGAGAAAGTGAGAAACATGCTGATAAAAGCTTAAGTTGTT 573
QY      154  ArgGluIleLeuArgValAlaLeuLysProGlyGlyIleLeuGlyValThrGluValValys 173
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      574 AATGAGTTGGACGAGGTTGACGCTCTGAGGACTAATATATATGATGATGGTGCAT 633
QY      174  ArgGluAlaGlyGlyGlyMetProValSerGlyAsp-----ArgTyrProThrGlyLeu 191
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      634 AGGAGTCTTGGT-----CCTTCAGAGAGAGATTGGAGCCATGGAGAAAAAGCTG 684
QY      192  -----ArgIleCysLeuAla-----Glu 197
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      685 CTAATATAGATATATGTAATAGCTTACTATTACCTGAGTGTGTTCTACTCTGCATATGTC 744
QY      198  GlnLeuLeuGlnSerLeuArgAlaAlaGlyPheGlnIleLeuAspTyrGluAspValSer 217
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      745 AAATCACTTCAAGTCCCTATCTCTCCAGGATATTAAGGCTGCTGATTTGG-----TCT 795
QY      218  SerArgThrArgTyrPheMetPro 225
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      796 GAGCGTGTTCCTCGTGTGGCCG 819

RESULT 5
CN008962      587 bp      mRNA      linear      EST 29-MAR-2004
LOCUS         WHE2647_B12_D23ZE wheat fusarium graminearum infected spike cDNA
DEFINITION    Library Triticum aestivum cDNA clone WHE2647_B12_D23, mRNA
sequence.
ACCESSION    CN008962
VERSION      CN008962.1  GI:45798994
KEYWORDS     EST.
SOURCE       Triticum aestivum (bread wheat)
ORGANISM     Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 587)
Anderson,O.D., Chao,S., Han,P.S., Heininen,S., Hsieh,C.C., Kang,Y.,
Kruger,W.M., Iazo,G.R., Miller,S., Muehlbauer,G.J., Miller,R.,
Pritsch,C., Rausch,C.J., Seaton,C.L., Tong,J.C., Vance,C. and
Wilson,C.F.
The structure and function of the expressed portion of the wheat
genomes - Fusarium graminearum infected spike cDNA library
Unpublished (2001)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20. No effort was taken
to identify ESTs of fungal origin from this library, thus this EST
could be of wheat or fungal origin.
Seq primer: T3 primer.
FEATURES
    source          1..587
                    /organism="Triticum aestivum"
                    /mol_type="mRNA"
                    /cultivar="Sumai3"
                    /db_xref="taxon:4565"
                    /clone="WHE2647_B12_D23"
                    /tissue_type="Spike"
                    /dev_stage="Adult plant"
                    /lab_host="E. coli SOLR"
                    /clone_lib="wheat Fusarium graminearum infected spike cDNA
                    library"
                    /note="Vector: Lambda Uni-ZAP XR, excised phagemid

```


QY 173 LysargLysAlaA---GlyGlyGlyMetProVal-----182
Db 471 ATGACCGACGCTTATGATACAAACACCCCGTCACCGTAGATCAGACTGTGGATCGAA 530
QY 183 SerGlyAspArgTrpProThrGlyLeuAglGlyCysLeuAlaGluGlnLeuGlnSer 202
Db 531 CAAAGGAGATGGAATTTTCGAACATGCTCAGG-----GCCAACGAGCCCTTGATGACC 581
QY 203 LeuArgAlaAlaGlyPheGluIleLeuAspTrpGluAspValSerArg 219
Db 582 TTTAAGCCCGCCGCTTTCGACTCATCAAGCTGAGACTTACAGACCTG 632
RESULT 7
LOCUS COI29317 824 bp mRNA linear EST 17-JUN-2004
DEFINITION GR_Eb26017.r GR_Eb Gossypium raimondii cDNA clone GR_Eb26017.3',
mRNA sequence.
ACCESSION COI29317 GI:48878295
VERSION COI29317.1 GI:48878295
KEYWORDS EST.
SOURCE Gossypium raimondii
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 824)
Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.,
TITLE Global assembly of Cotton ESTs
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
Plate: 26 row: O column: 17.
FEATURES
source
1.824
Location/Qualifiers
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eb26017"
/tissue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_lib="GR_Eb"
/note="Vector: PCMV.SPORT-6.1; Site 1: NotI; Site 2:
ECORV; Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Alignment Scores:

Pred. No.: 2e-15 length: 824
Score: 240.00 Matches: 69
Percent Similarity: 43.15% Conservative: 38
Best Local Similarity: 27.82% Mismatches: 107
Query Match: 16.62% Indels: 34
DB: 7 Gaps: 8

US-10-069-353A-8 (1-275) x COI29317 (1-824)

QY 3 ProGlyGlyAlaProThrSerGlnGlnValGlyGlnMetTrpAspLeuValThrProLeu 22
Db 35 CCCGATGCACGTCATTCGACGAGGAGATGCGGAGTTCTACGACCAAGTCGTCGGGCTT 94
QY 23 LeuAnserValAlaGlyGlyProCysAlaIleHisGlyTrpTrpGluAspGly 42
Db 95 TGGGAAGACATTTGGGGTGAC-----CATATGACATCATGTTTATACATCGGATTC 148

QY 43 ArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArgThrVal 62
Db 149 AATGTTTACAGGCTCCGATCATCATCAGCTGCCAGATCCGAATGATC---GAGAAATCGCTC 205
QY 63 LeuAspGlyGlyVal-----ArgLeuLeuAspValGly 73
Db 206 CGTTTCGCGTGAATATACCGATATCCAGTACCAAGCCCAAGCAATCATGATGTTTGA 265
QY 74 CysGlyThrGlyGlnProAlaLeuArgValAlaArgAspAsnAlaIleGlnIleThrGly 93
Db 266 TGTGGATAGGAGGAGGAGCTCTAGATCATTTAGCAAGAAATTTGGACGCAATGTCAGGC 325
QY 94 IlerhValSerGlnValGlnValAlaIleAlaAspCysAlaArgGluArgGlyLeu 113
Db 326 ATTACTTTGAGCCCTGTCGACAGCTCGAAGGGCTATGTTCTTGAAGATGAAGACGTA 385
QY 114 SerHisArgValAspPheSerCysValAspAlaMetSerLeuProTrpProAspAsnAla 133
Db 386 GCAGACAAAGCTTCATTTCAAGTTGCAGATGCTCTGAAGCAACCATTCCTGATGTCAG 445
QY 134 PheAspAlaAlaTrpAlaMetGlnSerLeuLeuMetSerGluProAspArgAlaIle 153
Db 446 TTTGATCTAGTTGGTCTATGGAAGTGAGACACATGCTGATTAAGCTTAAGTTGTT 505
QY 154 ArgGluIleLeuArgValLeuLeuSerProGlyGlyIleLeuGlyValThrGluValValys 173
Db 506 AATGAGTTGGACGACGATTCGACGCTCTCGAGGCATATATATATGATGACATGCGCAT 565
QY 174 ArgGluAlaGlyGlyGlyMetProValSerGlyAsp-----ArgTrpProThrGlyLeu 191
Db 566 AAGGATCTTGGT-----CCCTCAGAGAGATTTGAGCCATGAGCAAAAAAGCTG 616
QY 192 -----ArgIleCysLeuAla-----Glu 197
Db 617 CTAAATAGAGATATGTAATGCTTACTTACTGATGCTGTCTTACTTGCATATGTC 676
QY 198 GlnLeuLeuGlnSerLeuArgAlaAlaGlyPheGluIleLeuAspTrpGluAspValSer 217
Db 677 AAACACTTTCAGTCCCTATCTCTCCAGATATTAAGCTGCTGATTCG-----TCT 727
QY 218 SerArgThrArgTrpPheMetPro 225
Db 728 GAGCGTGTGGCCCGCTTTGGCCG 751
RESULT 8
LOCUS CF515375 902 bp mRNA linear EST 09-SEP-2003
DEFINITION CAP0001_IVF_C10 Vitis vinifera cv. cabernet sauvignon (Clone 8)
Petiole - CAP Vitis vinifera cDNA clone CAP0001_IVF_C10 5', mRNA
sequence.
ACCESSION CF515375
VERSION CF515375.1 GI:34547143
KEYWORDS EST.
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.
1 (bases 1 to 902)
Goed da Silva,F., Iandolo,A., Lim,H., Baek,J., Leslie,A., Xu,J.,
Jones,K. and Cook,D.
Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'
berries at various developmental stages
Unpublished (2003)
CONTACT: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcoc@ucdavis.edu
Seq primer: ACGATCCGACATATGCGC.

ORIGIN

Alignment Scores:

Pred. No.:	2,19e-15	Length:	669
Score:	238.50	Matches:	65
Percent Similarity:	46.02%	Conservative:	39
Best Local Similarity:	28.76%	Mismatches:	91
Query Match:	16.52%	Indels:	31
DB:	6	Gaps:	5

US-10-069-353a-8 (1-275) x CA220497 (1-669)

```

QY 9 SerGInGInValGlgYInMeTYrAsPLeuValThrProLeuLeuAsnSerValAlaGly 28
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 22 ACTGATATGTTAAACAAATATATATATATCTTGACACTAGCTCTATAGATAGTGGGGT 81
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 29 GLyProCySaAlaIleHshIGLyTYrTrpGluAsnAspGlyArGAlaSerTrpGlnGln 48
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 82 GAATCCTTCACATTGGCTCAC---AGATGGAATGGAGAA-----TCCTTAGCTGAA 129
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 49 AlaAlaAspArgLeuThrAspLeuValAlaGluArgThrValIleuAspGlyGlyValArg 68
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 130 AGCATCAACCGTCATGAGACATTTCTTCCCTCGACGCTTGGAACCAAGAAATGAAG 189
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 69 LeuLeuAspValGlyCySgLYThrgLYInProAlaLeuArgValAlaArgAspAsnAla 88
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 190 GTTTTGATGTGGGCTGCGGAATAGGTGGACCATGAGAAATTCAGAAATTTAGCTCA 249
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 89 IlegInIleThrgLYIleThrValSerGInValGInValAlaIleAlaIlaAspCySaAla 108
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 250 ACTTCAGTTACTGGGTGATGAATACATGAAATCCAGATTAACGAGGGAAGAGCTCAAT 309
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 109 ArgGluArgGlyLeuSerHisArgValAspPheSerCyValAspAlaMetSerLeuPro 128
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 310 CGGTGGCGAGGGATTAGTGAACCTGCCATTTTTCACAGCAGACTTATGAAGATGCCG 369
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 129 TyrProAspAsnAlaPheAspAlaAlaATrPalAMetGInSerLeuLeuMetSerGlu 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 370 TTTGATGATTAACACTTTTATGCTGTTATGCAATTAGAGCGCAATGTATCATCAGCTGAT 429
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 149 ProAspArgAlaIleArgGluIleLeuArgValIleuArgProGlyGlyIleLeuGlyVal 168
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 430 CCAAGTTGGCTGTATAAAGAGATATACCGTGTGTGAACCTGGCGCAAGTCTTCCTCTGA 489
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 169 ThrGlu-----ValIleAsArg-----ValIleAsArg----- 174
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 490 TATGAGTGTGATTAATCTATCATGTATCCAAACACGCAACCCATTAAGAGATTAAAG 549
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 175 -----GluAlaGlyGlyGlyMetProValSerGlyAspArgTrpProThrGlyLeu 191
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 550 GATGAATATTGAGCTTGGCAATGGCTGCCA----- 579
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 192 ArgGlyLeuValGluGlnLeuLeuGlnSerLeuArgAlaAlaGlyPheGluIleLeu 211
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 580 GATATCAAGAGCACTAAGCAATGTCTTCAGCAGATCAAGATGCTGTTTGAAGTGTGT 639
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 212 AspTrpGluAspValSer 217
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 640 GGAGATAAAGATCTAGCT 657
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 10
LOCUS AY103656 1665 bp mRNA linear HTC 16-OCT-2002
DEFINITION Zea mays PC0072714 mRNA sequence.
ACCESSION AY103656
VERSION AY103656.1 GI:21206734
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 1665)

```

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

1..1665

/organism="Zea mays"

/mol_type="mRNA"

/db_xref="maizedb:633975"

/db_xref="taxon:4577"

/clone_lib="Maize Mapping Project/Dupont Cornsensus Library"

/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

Alignment Scores:

Pred. No.:	7.64e-15	Length:	1665
Score:	238.50	Matches:	67
Percent Similarity:	46.19%	Conservative:	42
Best Local Similarity:	28.39%	Mismatches:	94
Query Match:	16.52%	Indels:	33
DB:	3	Gaps:	6

US-10-069-353a-8 (1-275) x AY103656 (1-1665)

```

QY 9 SerGInGInValGlgYInMeTYrAsPLeuValThrProLeuLeuAsnSerValAlaGly 28
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 382 ACTGATATGTTAATAAATACTATGATCTTGCACATGCTCTATAGATAGTGGGGT 441
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 29 GLyProCySaAlaIleHshIGLyTYrTrpGluAsnAspGlyArGAlaSerTrpGlnGln 48
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 GAATCCTTCACATTGGCTCAC---AGATGGAATGGAGAA-----TCCTTAGCTGAA 489
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 49 AlaAlaAspArgLeuThrAspLeuValAlaGluArgThrValIleuAspGlyGlyValArg 68
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 490 AGCATCAAGCAATGACCATTTTCTTCCCTCGCAACTGTTGTAACCAAGAAATGAAG 549
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 69 LeuLeuAspValGlyCySgLYThrgLYInProAlaLeuArgValAlaArgAspAsnAla 88
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 550 GTTTTGAATGTGGCTGCGAATAGGTGGACCACTGAGAAATTCAGAAATTTAGCTCA 609
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 89 IlegInIleThrgLYIleThrValSerGInValGInValAlaIleAlaIlaAspCySaAla 108
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 610 ACTTCAGTTACCGGATTCATTAACACAGAAATACAGATTAACGAGGGAAGAGCTCAAC 669
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 109 ArgGluArgGlyLeuSerHisArgValAspPheSerCyValAspAlaMetSerLeuPro 128
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 670 CGTTTAGCGAGAAATTAGGGAACATGTATTTTCTCAAGCGGAGCTTCATTAAGATGCCG 729
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 129 TyrProAspAsnAlaPheAspAlaAlaATrPalAMetGInSerLeuLeuMetSerGlu 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 730 TTCATGCAACACTTTTATGATGCTGTATACCCATTAGGCAACATGTATGACACCTGAT 789
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 149 ProAspArgAlaIleArgGluIleLeuArgValIleuArgProGlyGlyIleLeuGlyVal 168
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```
/lab_host="DH10B"  
/clone_lib="GR_Eb"  
/note="Vector: PCWV.SPORT-6.1; Site 1: NotI; Site 2:  
EcoRV; Library made by invitrogen with RNA supplied by  
Wendle lab. Directional cloned into NotI-EV. Colonies  
plated/picked by AGI. More glycerol clones held in -80."
```

ORIGIN

Alignment Scores:

Pred. No.:	4,51e-15	Length:	786
Score:	236.50	Matches:	65
Percent Similarity:	43.70%	Conservative:	39
Best Local Similarity:	27.31%	Mismatches:	103
Query Match:	16.38%	Indels:	31
DB:	7	Gaps:	7

US-10-069-353A-8 (1-275) x COI20890 (1-786)

```
QY 3 ProGlyGlyAlaProThrsSerGlnGlnValGlyGlnMetCysAspLeuValThrProLeu 22  
DB 81 CCCGATGACAGTCGATGACAGAGGATAGCGAGTTCTACGACCGATCGTCTGGGCTT 140  
QY 23 LeuAenSerValAlaGlyGlyProCysAlaIleHisGlyTyrTrpGlnuSnpGly 42  
DB 141 TGGGAAGACATTTGGGGTGAC-----CATATGCATCATGTTTATCGATCCGAGATTCC 194  
QY 43 ArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGlnuArgThrVal 62  
DB 195 AATGTTTACGGCTCCGATCATCAGCTGCCAGATCCGAAATGATC---GAAAGATCGCTC 251  
QY 63 LeuAspGlyGlyVal-----ArgLeuLeuAspValGly 73  
DB 252 CGTTTCGCTGGAATACCGATATCCAGCAAGACCCCAAGACATATGTTGATTTGGA 311  
QY 74 CysGlyThrGlyGlnProAlaLeuArgValAlaArgAspSnaAlaIleGlnIleThrGly 93  
DB 312 TGTGGATAGGAGGACGCTCTAGATATTAGCAAGAAATTTGGACGCAATGTCAAGGC 371  
QY 94 IleThrValSerGlnValGlnValAlaIleAlaAspCysAlaArgGlnuArgGlyLeu 113  
DB 372 ATTACTTTGAGCCCTGCTCCAGCTGGAAGGGCTAATGTTCTTGCATTAAGATGAAGACTA 431  
QY 114 SerHisArgValAspPheSerCysValAspAlaMetSerLeuProTyrProAspSnaAla 133  
DB 432 GAGACAGAGTTTCAATTCAGTTCAGATGCTCGAAGCAACATTCCTCATGTGTCAG 491  
QY 134 PheAspAlaAlaTrpAlaMetGlnSerLeuLeuGlnMetSerGlnProAspArgAlaIle 153  
DB 492 TTGTATCTAGTTGGTCTATGGAAGGTGAGAACACATGCTGATTAAGCTAATGTTGTT 551  
QY 154 ArgGlnIleLeuArgValLeuLeuSnpProGlyGlyIleLeuGlyValThrGlnuValuLys 173  
DB 552 AATGATTTGGACGAGATTGACGCTCTGGAGGCACATTAATATATGACATGCGCAT 611  
QY 174 ArgGlnuAlaGlyGlyGlyMetProValSerGlyAsp-----ArgTyrProThrGlyLeu 191  
DB 612 AGGGATCTTGGT-----CCCTCAGAGAGGATTTGAGCCATGGAGAAAGCTG 662  
QY 192 -----ArgIleCysLeuAla-----Gln 197  
DB 663 CTAAATAGAGATATGTAATGCTTACTTATTACCTGAGTGGTGTCTTACTTGCATATGTC 722  
QY 198 GlnLeuLeuGlnuSerLeuArgAlaAlaGlyPheGlnIleLeuAspTrpGlnuAsp 215  
DB 723 AAACACTTCACTCCCTATCTCTCCAGAGATATAAAGGCTGCTGATTTGGTCTGAG 776
```

RESULT 13

CO009335/c

LOCUS

DEFINITION

ACCESSION

VERSION

CO009335 961 bp mRNA linear EST 09-JUN-2004
EST9797670 Coccidioides posadasii spherule cDNA library, 0.4 to 2.3
kb Coccidioides posadasii cDNA clone CIBB26 3' end, mRNA sequence.
CO009335
VERSION CO009335.1 GI:48516224

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Other ESTs: EST97671

Contact: Gardner MJ

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301 838 3519

Fax: 301 838 0208

Email: gardner@tigr.org

Location/Qualifiers

1..961

/organism="Coccidioides posadasii"

/mol_type="mRNA"

/strain="C735"

/db_xref="taxon:199306"

/clone="CIBB26"

/dev_stage="spherules"

/lab_host="E. coli DH10B, T1 phage resistant"

/clone_lib="Coccidioides posadasii spherule cDNA library, 0.4 to 2.3 kb"

/note="Vector: pExpress 1; Site 1: Not I; Site 2: Eco RV; Coccidioides posadasii spherule cDNA library, 0.4 to 2.3 kb"

ORIGIN

Alignment Scores:

Pred. No.:	5.94e-15	Length:	961
Score:	236.50	Matches:	67
Percent Similarity:	47.93%	Conservative:	37
Best Local Similarity:	30.88%	Mismatches:	90
Query Match:	16.38%	Indels:	23
DB:	7	Gaps:	5

US-10-069-353A-8 (1-275) x CO009335 (1-961)

```
QY 16 TyrAspLeuValThrProLeuLeuAsnSerValAlaGlyGlyPro-----CysAla 32  
DB 951 TCAATCTGGCGGAGGACATGATGAGATAGAGGGCGGCTTCTTCCACTTGTGCT 892  
QY 33 IleHisGlyTyrTrpGlnuAspGlyArgAlaSerTrpGlnGlnAlaAlaAspArg 52  
DB 891 TTCGCGTATGCG-----GAACTTTCCGTCAGGCTATGTCGA 853  
QY 53 LeuThrAspLeuValAlaGlnuArgThrValLeuAspGlyGlyValArgLeuLeuAspVal 72  
DB 852 CACGACATTAATCTCGGACACTCATCGGCTCCAGAGAGGCGCATGTGCGGAGAAAGT 793  
QY 73 GlyCysGlyThrGlyGlnProAlaLeuArgValAlaArgAspSnaAlaIleGlnIleThr 92  
DB 792 GCGTCGGCGCTTGGTGGCCCCCGCAGAGAAATTCCAAATTTGGCGGTCGAATTTGTT 733  
QY 93 GlyIleThrValSerGlnValGlnValAlaIleAlaAlaAspCysAlaArgGlnuArgGly 112  
DB 732 GGTTAATATAACAACGACTATCAAAATGAAACGCCCACTCGATATGCGGAGAAAGT 673  
QY 113 LeuSerHisArgValAspPheSerCysValAspAlaMetSerLeuProTyrProAspSna 132  
DB 672 TTGTCAAAGAGCTCCCTTCACAAAGGTGACTTATGCAAAATGCTTTGAACAAAT 613  
QY 133 AlaPheAspAlaAlaTrpAlaMetGlnSerLeuLeuGlnuMetSerGlnuProAspArgAla 152  
DB 612 ACATTTATGATGATATGCAATGCAAGGCTACAGCTACACCTCTTGGCTTGAAGGCGTC 553  
QY 153 IleArgGlnIleLeuArgValLeuLeuSnpProGlyGlyIleLeuGlyValThrGlnuVal 172
```

```

Db      552 TACAGACAGATTACACTATTAAACCCGGCGGTGTTTGTGTATACGAATGGC 493
Qy      173 LysargluAa---GlyglyMetProVal----- 182
Db      492 ATGACCGAGCTTATGATATACAAACACCCGTCACCGTAGATCAGACTTGATACGA 433
Qy      183 SerGlyAspArgTrpProthrglyLeuAarglyCysLeuAlaGluInleuLeuSer 202
Db      432 CAAGGAGATGGAATTTGGAACATGTCAGG-----GCACCCGAGCCCTTGATGCC 382
Qy      203 LeuArgAlaAlaGlyPheGluLeuAspTrpGluAspValSerArg 219
Db      381 TTTAAGCGCCGCTTTCAGACTCATCAAGCTGAGAGCTTACACACCT 331

RESULT 14
AJ431900      634 bp      mRNA      linear      EST 15-MAR-2002
LOCUS      AJ431900      500007      Hordeum vulgare cDNA clone 80000700049603FL, mRNA
DEFINITION      sequence.
ACCESSION      AJ431900      GI:19520352
VERSION      AJ431900.1
KEYWORDS      EST.
SOURCE      Hordeum vulgare
ORGANISM      Hordeum vulgare
REFERENCE      Saren,A.-M., Tanskanen,J., Paulin,L. and Schulman,A.H.
AUTHORS      Barley EST's
TITLE      Unpublished (2002)
JOURNAL      Contact: Schulman AH
COMMENT      Institute of Biotechnology
University of Helsinki
P.O.Box 56 (Wilkinlaari 6A), University of Helsinki FIN-00014,
Finland.
FEATURES
     source          1..634
                     /organism="Hordeum vulgare"
                     /mol_type="mRNA"
                     /db_xref="taxon:4513"
                     /clone="S0000700049603FL"
                     /dev_stage="shoot"
                     /clone_lib="S00007"
                     /note="2,-3,-4-days after germination"

ORIGIN
Alignment Scores:
Pred. No.:      3,81e-15      Length:      634
Score:      236.00      Matches:      65
Percent Similarity:      45.37%      Conservative:      33
Best Local Similarity:      30.09%      Mismatches:      86
Query Match:      16.34%      Indels:      32
DB:      1      Gaps:      6

US-10-069-353a-8 (1-275) x AJ431900 (1-634)
Qy      16 TyAspLeuValThrProLeuLeuAspSerValAlaGlyGlyProCysAlaIleHis 35
Db      8 TATGATCTTCTACTACTCTACAGACTATGCGGCGTGAATCTTCCACTTGCTCAC 67
Qy      36 GlyTrpTrpGluAspGlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAsp 55
Db      68 ---AGGTGAGAGTGAAGAA-----TCTTTAAGGAGAAATTCACCAACGCGACGACAC 115
Qy      56 LeuValAlaGluArgThrValLeuAspGlyGlyValArgLeuAspValGlyCysGly 75
Db      116 TTTCTGCGCTTACGAGCTTGAGTTAAACGAGGATGAAGTTTGGATGTCGGCTGGGA 175
Qy      76 ThrGlyGlnProAlaLeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThr 95
Db      176 ATAGCGCGCGCTTAAAGAAATTGCGAGATTAGCTCCACCTCAGTTACTCGATTGAC 235

```

```

Qy      96 ValSerGlnValGlnValAlaIleAlaAspCysAlaArgGluArgGlyLeuSerHis 115
Db      236 AACACGACTACCAATATACAGGGAAGAGACTTATCGGTCTGTAGACTTGATGCA 295
Qy      116 ArgValAspPheSerCysValAspAlaMetSerLeuProTrpProAspAsnAlaPheAsp 135
Db      296 ACTTGATATTTGTCAAGGACAGACTTCATGAAGATGCAATCTCTGATTAACACTTTGAT 355
Qy      136 AlaAlaTrpAlaMetGlnSerLeuLeuGluMetSerGluProAspArgAlaIleArgGlu 155
Db      356 GCTGTTTATGCAATGAGGACATGCCAGCGCTGATCCGCTGCTCTATTAAGAG 415
Qy      156 IleLeuArgValLeuLeuProGlyGlyIleLeuGlyValThrGlu----- 170
Db      416 ATCTACCGTATTTAAACCCGGGACAGTGTTCGTGATATGATGATGATGATACCAT 475
Qy      171 -----ValValLysArg-----GluAlaGlyGly 178
Db      476 CACTATGATCCAAACATGACCAACCAAGAGATTAAAGATGAATTGAGCTGTGATAC 535
Qy      179 GlyMetProValSerGlyAspArgTrpProthrglyLeuAarglyCysLeuAlaGluIn 198
Db      536 GGCTCTGCCA-----GATATCAAGAACTACTGGGCA 565
Qy      199 LeuLeuGluSerLeuArgAlaAlaGlyPheGluLeuAspTrpGlu 214
Db      566 TGTCTTCAAGCTTTAAAGATCTGGGTTTGAGCTTAT--TGGGAT 610

RESULT 15
BJ271067      650 bp      mRNA      linear      EST 09-APR-2002
LOCUS      BJ271067      650007      Ogiwara unpublished cDNA library, Wh_oh Triticum
DEFINITION      aestivum cDNA clone whon23p14 5', mRNA sequence.
ACCESSION      BJ271067      GI:20096630
VERSION      BJ271067.1
KEYWORDS      EST.
SOURCE      Triticum aestivum (bread wheat)
ORGANISM      Triticum aestivum
REFERENCE      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 650)
Ogiwara,Y. and Murai,K.
TITLE      Expressed genes in Triticum aestivum
JOURNAL      Unpublished (2002)
COMMENT      Contact: Tadasi Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
     source          1..650
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="whon23p14"
                     /tissue_type="pistil at heading date"
                     /dev_stage="Feekes' scale 10.5"
                     /clone_lib="Y. Ogiwara unpublished cDNA library, Wh_oh"

ORIGIN
Alignment Scores:
Pred. No.:      6.51e-15      Length:      650
Score:      234.00      Matches:      66
Percent Similarity:      44.05%      Conservative:      34
Best Local Similarity:      29.07%      Mismatches:      95
Query Match:      16.20%      Indels:      32
DB:      4      Gaps:      6

US-10-069-353a-8 (1-275) x BJ271067 (1-650)

```

```
OY      5 GlyAlaProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThrProLeuLeuSn 24
      1 GGAAGAGAACTACACCGATATGTTAATAATACTATGATCTTGCTACTGCTTCTATGAG 60
OY      25 SerValAlaGlyGlyProCysAlaIleHisIleGlyTyrTrpGluSnAspGlyArgAla 44
      61 TATGCTGGGTGAGTCCCTTCCACTTTCCTCAC--AGGTGGAAATGAGAA----- 108
OY      45 SerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArgThrValLeuAsp 64
      109 TCTTTAAGGGAAGACATCAACGCGACGCACTTTCGCTTTCACAGCTTGAGTTGAAA 168
OY      65 GlyGlyValArgLeuLeuAspValGlyCysGlyThrGlyGlnProAlaLeuArgValAla 84
      169 CCAAGGATGAAGATTGATGATCGGCTGTGAAATAGCGCGCCATTAAGAGAAATTGGG 228
OY      85 ArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGlnValAlaIleAla 104
      229 AGATTAGCTCCACTCAGCTTACTGATTTGAACAACAGACTACAGATACTAGGGGA 288
OY      105 AlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSerCysValAspAla 124
      289 AAGGCGCTTAATCGGTCGGTAGGACTTGCGCACTTGATTTGTCAAGGCAAGACTTC 348
OY      125 MetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMetGlnSerLeuLeu 144
      349 ATGAAGATGCCATTCTCGATACACTTTTGATGCTGTTATGCCATCGAGGCAACATGC 408
OY      145 GluMetSerGluProAspArgAlaIleArgGluIleLeuArgValLeuLysProGlyGly 164
      409 CACGCACCTGATCCGTTGGCTGCTACAAAGAGATCTACCGGTATTAAACCCGGGCA 468
OY      165 IleLeuGlyValThrGlu-----ValVal 172
      469 TCTTTGCTGTATGATGAGTGCATTTACCGATCACTATGATCCAAACATGCAACCCAC 528
OY      173 LysArg-----GluAlaGlyGlyMetProValSerGlyAspArgTrp 187
      529 AAGAGGATTAAGATGAATTAAGCTTGGGATGCTGCCA----- 570
OY      188 ProThrGlyLeuArgIleCysLeuAlaGluGlnLeuLeuGluSerLeuArgAlaAlaGly 207
      571 -----GATATCAGAACTACGCGCAATGCTTCAAGCTTAAGATGCTGGG 618
OY      208 PheGluIleLeuAspTrpGlu 214
      619 TTTGAGTTATT--TGGGAC 636
Db
```

Search completed: August 29, 2005, 22:18:49
Job time : 3551 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 29, 2005, 19:39:44 ; Search time 228 Seconds
(without alignments)
1973.579 Million cell updates/sec

Title: US-10-069-353a-8

Perfect score: 1444
Sequence: 1 VLEGGAPTSQQVGMVDTLT.....YERKADMGVAILTANKPVG 275

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE:frame+p2n.model -DEV:xlh
-O:/cgnt2_1/uspro.spool/us10069353/runat.26082005.172054.14693/app.query.fasta_1.455
-DB:issued_patents.na -OPMT:fastcap -SUFFIX:mi -MINMATCH=0.1 -LOCPCL=0
-LOOPEXT=0 -UNITS:bits -START=1 -END=1 -MATRIX:dlosum62 -TRANS:human40.cdt
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US10069353 @CGN 1.1 69 @runat.26082005.172054.14693 -NCPUL=6 -ICPU=3
-NO MMAP -LARGEDUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

- 1: /cgnt2_6/ptodata/1/ina/5A.COMB.seq:*
- 2: /cgnt2_6/ptodata/1/ina/5B.COMB.seq:*
- 3: /cgnt2_6/ptodata/1/ina/6A.COMB.seq:*
- 4: /cgnt2_6/ptodata/1/ina/6B.COMB.seq:*
- 5: /cgnt2_6/ptodata/1/ina/PTUS.COMB.seq:*
- 6: /cgnt2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1444	100.0	80161	3	US-09-036-987A-1
2	1444	100.0	80161	3	US-09-370-700-1
3	1444	100.0	80161	4	US-09-603-207-1
4	435.5	30.2	80161	3	US-09-036-987A-1
5	435.5	30.2	80161	3	US-09-370-700-1
6	435.5	30.2	80161	4	US-09-603-207-1
7	379	26.2	6085	3	US-09-029-603-4
8	376.5	26.1	852	4	US-09-266-965-36
9	376.5	26.1	53500	4	US-09-266-965-76
10	345	23.9	828	4	US-09-266-965-35
11	320.5	22.2	2122	3	US-09-029-603-1
12	241.5	16.7	1444	4	US-09-885-723-21

13	216	15.0	1164	4	US-09-248-796A-3873	Sequence 3873, Ap
14	216	15.0	1400	3	US-09-041-718-1	Sequence 1, Appli
15	210.5	14.6	1990	4	US-09-118-637A-3	Sequence 3, Appli
16	206	14.3	954	4	US-09-118-637A-1	Sequence 1, Appli
17	200.5	13.9	1421	1	US-09-885-723-22	Sequence 22, Appli
18	198	13.7	441529	3	US-09-103-840A-1	Sequence 1, Appli
19	196	13.6	957	3	US-09-382-906A-1	Sequence 1, Appli
20	195.5	13.5	1152	4	US-09-614-221A-523	Sequence 523, App
21	193.5	13.4	4403765	3	US-09-103-840A-2	Sequence 2, Appli
22	187.5	13.0	1369	4	US-09-439-554-29	Sequence 29, Appli
23	181.5	12.6	1062	4	US-09-902-540-3096	Sequence 3096, Ap
24	181.5	12.6	15499	4	US-09-902-540-1140	Sequence 1140, Ap
25	179.5	12.4	759	4	US-09-583-110-637	Sequence 637, App
26	179.5	12.4	780	4	US-09-107-433-1646	Sequence 1646, Ap
27	179.5	12.4	1249	4	US-09-885-723-20	Sequence 20, Appli
28	177.5	12.3	3236	3	US-08-961-527-1022	Sequence 222, App
29	170.5	11.8	9377	3	US-09-221-017B-1002	Sequence 1002, Ap
30	168.5	11.7	1131	4	US-09-252-991A-2322	Sequence 2322, Ap
31	168.5	11.7	1215	4	US-09-252-991A-2513	Sequence 2513, Ap
32	168.5	11.7	1215	4	US-09-252-991A-2595	Sequence 2595, Ap
33	162.5	11.3	783	4	US-09-489-038A-5480	Sequence 5480, Ap
34	162	11.2	6599	4	US-09-902-540-724	Sequence 724, App
35	161	11.1	597	4	US-09-252-991A-16384	Sequence 16384, A
36	161	11.1	1866	4	US-09-252-991A-16213	Sequence 16213, A
37	154	10.7	1830121	4	US-09-557-884-1	Sequence 1, Appli
38	154	10.7	1830121	4	US-09-643-990A-1	Sequence 1, Appli
39	153.5	10.6	696	4	US-09-902-540-3923	Sequence 3923, Ap
40	153.5	10.6	10263	4	US-09-902-540-1024	Sequence 1024, Ap
41	152.5	10.6	747	4	US-09-902-540-9423	Sequence 9423, Ap
42	152.5	10.6	11358	4	US-09-902-540-1075	Sequence 1075, Ap
43	152	10.5	20303	4	US-09-902-540-1183	Sequence 1183, Ap
44	151.5	10.5	4403765	3	US-09-103-840A-2	Sequence 2, Appli
45	151.5	10.5	4411529	3	US-09-103-840A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-036-987A-1
Sequence 1, Application US/09036987A
Patent No. 6143526
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Maddux, Krishnamurthy
APPLICANT: Merlo, Donald J.
APPLICANT: Treadway, Patti J.
APPLICANT: Turner, Jan R.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: BioSynthetic Genes For Spinosyn Insecticide
TITLE OF INVENTION: Production
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow AgroSciences LLC Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,987A
FILING DATE: 09-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608


```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816
; TELEFAX: (317)337-4847
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80161 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-036-987A-1

Alignment Scores:
Pred. No.: 4,54e-160 Length: 80161
Score: 1444.00 Matches: 275
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-069-353A-8 (1-275) x US-09-036-987A-1 (1-80161)
QY 1 ValLeuProGlyGlyAlaProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThr 20
Db 20168 GTGTTCAGGTCGTCGACCAATCGCAGGTTGGGCAATGATGACCTGGTCAGC 20227
QY 21 ProLeuLeuAsnSerValAlaGlyGlyProCysAlaIleHisGlyTyrTrpGluAsn 40
Db 20228 CCGTTGCTGAACCTCGGTCCGGGCGGCCCCCTGCGCATCACACCGCTACTGGAGAAC 20287
QY 41 AspGlyArgAlaSerTrpGlnGlnAlaAspArgLeuThrAspLeuValAlaGluArg 60
Db 20288 GACGGGCGGCGCTTCCTGCGCAGCAGCGCCGACCGGCTCACGACCTTGCCCGAACCG 20347
QY 61 ThrValLeuAspGlyGlyValArgLeuAspValGlyCysGlyThrGlyGlnProAla 80
Db 20348 ACCGAGCTCGATGGCGCGCTGACCTGCGATGTTGGGGTGGCGTACCGGACCAACCGG 20407
QY 81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGln 100
Db 20408 CTGGCGGTCGGCGCGGACCAACCGCATCCGATCACCGGATCACCGTCCAGCGAGTGCA 20467
QY 101 ValAlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120
Db 20468 GTGGCGATCGCGCTGATGTCGCGACGCGAACCTTAAGCCACCGGGTGGACTTTCG 20527
QY 121 CysValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
Db 20528 TCGTCGATGCATGTCCTGCGCGTACCGGACAAATGCTTTCGACGCGCGCTGGGCGCATG 20587
QY 141 GlnSerLeuLeuGlnMetSerGluProAspArgAlaIleArgGluIleLeuArgValLeu 160
Db 20588 CAGTCGCTGTCGAGATGTCGACCGGACCGTGCATCCGGGAATCTTTCGATGATCTC 20647
QY 161 LysProGlyGlyIleLeuGlyValThrGluValValValArgGluAlaGlyGlyMet 180
Db 20648 AAACCCGGGTGCATCTCGCGCTCACCGGATCGTCAACCGAAGCGGGCGCGGAGTG 20707
QY 181 ProValSerGlyAspArgTrpProThrGlyLeuArgIleCysLeuAlaGluGlnLeu 200
Db 20708 CCGGTCGTCGGGGAACGTTGGCGCGACCGGCTTCGGAATGCTTTCGATGAGAACCTTTCG 20767
QY 201 GlnSerLeuArgAlaAlaGlyPheGluIleLeuAspTrpGluAspValSerSerArgThr 220
Db 20768 GAATCGCTGCGTCGACGCGGAGTTTCGAGATCCTCGATTGGAGAGACGTGTCGAGGACC 20827
QY 221 ArgTyrPheMetProGlnPheAlaGluGluLeuAlaAlaHisGlnHisGlyIleAlaAsp 240
Db 20828 CGGTACTTCATGCCCCAGTTCCCGAAGAGCTGCTGGCGCACACACGAGGATCCCGGAC 20887
QY 241 ArgTyrGlyProAlaValAlaGlyTyrTrpAlaAlaValCysAspTrpGluSerTyrAla 260
Db 20888 AGGTACGGGCGCGCTGTCGCGGCTGGCGCGCGGCTGCGATTATGAGAAATATGCG 20947
```

```

QY 261 HisAspMetGlyTyrAlaIleLeuThrAlaArgLysProValGly 275
Db 20948 CACGACATGGGCTATGCAATTCGACGCGCGGAGACCGGTCGCGC 20992

RESULT 2
US-09-370-700-1
; Sequence 1, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madhuri, Krishnamurthy
; APPLICANT: Treadway, Patci J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/370,700
; EARLIER FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/36987
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 80161
; TYPE: DNA
; ORGANISM: Saccharopolyspora spinosa
US-09-370-700-1

Alignment Scores:
Pred. No.: 4,54e-160 Length: 80161
Score: 1444.00 Matches: 275
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-069-353A-8 (1-275) x US-09-370-700-1 (1-80161)
QY 1 ValLeuProGlyGlyAlaProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThr 20
Db 20168 GTGTTCAGGTCGTCGACCAATCGCAGGTTGGGCAATGATGACCTGGTCAGC 20227
QY 21 ProLeuLeuAsnSerValAlaGlyGlyProCysAlaIleHisGlyTyrTrpGluAsn 40
Db 20228 CCGTTGCTGAACCTCGGTCCGGGCGGCCCCCTGCGCATCACACCGCTACTGGAGAAC 20287
QY 41 AspGlyArgAlaSerTrpGlnGlnAlaAspArgLeuThrAspLeuValAlaGluArg 60
Db 20288 GACGGGCGGCGCTTCCTGCGCAGCAGCGCGACCGGCTCACGACCTTGTGCCGAACCG 20347
QY 61 ThrValLeuAspGlyGlyValArgLeuAspValGlyCysGlyThrGlyGlnProAla 80
Db 20348 ACCGAGCTCGATGGCGCGGTTTCGACTGCTCATGTGGGGTCCGGTACCGGACCAACGCG 20407
QY 81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGln 100
Db 20408 CTGGCGGTCGGCGCGGACCAACCGCATCCGATCACCGGATCACCGTTCAGCGAGTGCA 20467
QY 101 ValAlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120
Db 20468 GTGGCGATCGCGCTGATGTCGCGACGCGAAGCGGACTTAAGCCACCGGGTGGACTTTCG 20527
QY 121 CysValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
Db 20528 TCGTCGATGTCATGCTCTGCGCGTACCGGACAAATCTTTCGACGCGCGCTGGGCGCATG 20587
QY 141 GlnSerLeuLeuGlnMetSerGluProAspArgAlaIleArgGluIleLeuArgValLeu 160
Db 20588 CAGTCGCTGTCGAGATGTCGAAACCGGACCGTGCATCCGGGAATTCCTTTCGATCTC 20647
```



```
/ FILING DATE: 09-MAR-1998
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Stuart, Donald R
/ REGISTRATION NUMBER: 28,479
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (317)337-4816
/ TELEFAX: (317)337-4847
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 80161 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
US-09-036-987A-1

Alignment Scores:
Pred. No.: 1,86e-39 Length: 80161
Score: 435.50 Matches: 108
Percent Similarity: 51.67% Conservative: 47
Best Local Similarity: 36.00% Mismatches: 96
Query Match: 30.16% Indels: 49
DB: Gaps: 11

US-10-069-353A-8 (1-275) x US-09-036-987A-1 (1-80161)
QY 7 ProthrserGlnGlnValGlyGlnMetYrAspLeuValThrProleuLeuAnsSerVal 26
DB 13562 CCGACCGCGGATCAGATGAGTGCATCTTCGATGCGTGTG-----GCG 13521
QY 27 AlaGlyGlyProCysAlaIleHisGlyTyrTrp-----GluAnsAsp 41
DB 13520 CACGGGCGTCCC-----CTGCACACCGGTTACTGGGGCGGGCGGATCGGAGATGCC 13467
QY 42 GlyArgAlaSerTyrPngInAlaIleAspArgLeuThrAspLeuValAlaGluArgThr 61
DB 13466 GGTGCACACCGGTGCGATGCTGCGCCGACCAATGACCGACTGTTCATGCACAAAGGCC 13407
QY 62 ValLeuAspGlyGlyValArgLeuLeuAspValGlyCysGlyTyrGlyGlnProAlaLeu 81
DB 13406 GCGCTCCGTCGCGAGCGCACCTGTTGCACCTGGCGTGGCGCAATGGCGACCGCGTAGTC 13347
QY 82 ArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGlnVal 101
DB 13346 CGTGGCGCATGGCGCCAGCGCGCTTCGACGACCGGAATACCGTGAAGCCCAAGATCTC 13287
QY 102 AlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSerCys 121
DB 13286 GCGCGCGCACACGAGCTGCCCAACGAGACCGGACTGGCGCGGACGTTGAGTTGATGTA 13227
QY 122 ValAspAlaMetSerLeuProTyrProAlaAspAlaIleThrValMetGln 141
DB 13226 GTCGACGGCGCCAGCGCATGCTCCACCGGACGCTTTCTTCACGCGCGCATGGCGATGCA 13167
QY 142 SerLeuLeuGluMetSerGluProAspArgAlaIleArgGlnIleLeuArgValLeuLys 161
DB 13166 TCCGTCGTCGACATGTCGACCAAGCGCGCGGATCCGAGAGTCCACCGAATCTCGGA 13107
QY 162 ProGlyGly-----IleLeuGly-----ValThrGluValValLysArgGlnAlaGly 177
DB 13106 CCGCGCGCGCGGTGCTGCTCGGAGACATCATCATCTCGGTTGCACTCCCGCAAGATAC 13047
QY 178 GlyGlyMetProValSerGlyAspArgTyrProThrGlyLeuArgIleCysLeuAlaGlu 197
DB 13046 GCGGCGGCT-----TGG-----ACGGGACACGACCGCCATACCTTGAC 13008
QY 198 GlnLeuLeuGluSerLeuArgAlaIleGlyPheGlnIleLeuAspTyrGlnAspValSer 217
DB 13007 AGCTTACGCGCGCTGTCAGCGAAGCGGGTTCGAGATTCCTGAAGTCAACGACCTTCACG 12948
QY 218 SerArgThrArgTyrPheMetProGlnPheAlaGluGluLeu----- 231
```

```
DB 12947 GCACAGACCAAGTGCATGTCCTCTGATCTCCGACGAGTTCGCGAACTCGATGAG 12888
QY 232 -----AlaIleHisGlnHis-----GlyIle 238
DB 12887 CTCGCCGCGTCGACGCTCGCGCTGTGCGCACCTTACGACCAAGCTACTTGGAGACATC 12828
QY 239 AlaAspArgTyrGlyProAlaValAlaGlyTyrAlaAlaValAlaCysAspTyrGlnLys 258
DB 12827 GCGGGGAAGCAAGGACCGCGGACGACGACGATGCGCGCGTTGGCGGATACCGGAA 12768
QY 259 -----TyrAlaHisAsp-----MetGlyTyrAlaIleLeuThrAlaArgLys 272
DB 12767 CATCCGATTTACGCGCAAGAGGAAAGCATGGGTTTCATGCTCTCGACGCTCGAAG 12708

RESULT 5
US-09-370-700-1/c
/ Sequence 1, Application US/09370700
/ Patent No. 6274350
/ GENERAL INFORMATION:
/ APPLICANT: Balitz, Richard H
/ APPLICANT: Broughton, Mary C
/ APPLICANT: Crawford, Kathryn P
/ APPLICANT: Madduri, Krishnamurthy
/ APPLICANT: Treadway, Patsi J
/ APPLICANT: Turner, Jan R
/ APPLICANT: Waldron, Clive
/ TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
/ FILE REFERENCE: 50489 DIV1
/ CURRENT APPLICATION NUMBER: US/09/370,700
/ CURRENT FILING DATE: 1999-08-09
/ EARLIER APPLICATION NUMBER: US 09/36987
/ NUMBER OF SEQ ID NOS: 39
/ SOFTWARE: Patent In Ver. 2.0
/ SEQ ID NO 1
/ LENGTH: 80161
/ TYPE: DNA
/ ORGANISM: Saccharopolyspora spinosa
US-09-370-700-1

Alignment Scores:
Pred. No.: 1,86e-39 Length: 80161
Score: 435.50 Matches: 108
Percent Similarity: 51.67% Conservative: 47
Best Local Similarity: 36.00% Mismatches: 96
Query Match: 30.16% Indels: 49
DB: Gaps: 11

US-10-069-353A-8 (1-275) x US-09-370-700-1 (1-80161)
QY 7 ProthrserGlnGlnValGlyGlnMetYrAspLeuValThrProleuLeuAnsSerVal 26
DB 13562 CCGACCGCGGATCAGATGAGTGCATCTTCGATGCGTGTG-----GCG 13521
QY 27 AlaGlyGlyProCysAlaIleHisGlyTyrTrp-----GluAnsAsp 41
DB 13520 CACGGGCGTCCC-----CTGCACACCGGTTACTGGGGCGGGCGGATCGGAGATGCC 13467
QY 42 GlyArgAlaSerTyrPngInAlaIleAspArgLeuThrAspLeuValAlaGluArgThr 61
DB 13466 GGTGCACACCGGTGCGATGCTGCGCCGACCAATGACCGACTGTTCATGCACAAAGGCC 13407
QY 62 ValLeuAspGlyGlyValArgLeuLeuAspValGlyCysGlyTyrGlyGlnProAlaLeu 81
DB 13406 GCGCTCCGTCGCGAGCGCACCTGTTGCACCTGGCGTGGCGCAATGGCGACCGCGTAGTC 13347
QY 82 ArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGlnVal 101
DB 13346 CGTGGCGCATGGCGCCAGCGCGCTTCGACGACCGGAATACCGTGAAGCCCAAGATCTC 13287
QY 102 AlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSerCys 121
```

```
Dh 13286 GCCCGCCGACGAGCTCGCAGAACGAGACCGGAGCTGCGGAGCTTGATGATCTTA 13227
Qy 122 ValAaPAlaMeSerLeuProTyrProAspAaAlaPheAspAlaAlaTTPAlaMeGln 141
Dh 13226 GTCGAGCGGCGCCGAGCTGCGCTTACCGGAGCTTTCTTTCAGGCGCCGATGGGAGATGACG 13167
Qy 142 SerLeuGluMeSerGluProAspAaAlaAlaArgGluIleLeuArgValLeuLys 161
Dh 13166 TCCGTGTGCAGATCTGGACGACGCGCCGCGATCGGAGGTCCAGCAGATCTCTGGA 13107
Qy 162 ProGlyGly-----IleuGly-----ValThrGluValValLysArgGluAlaGly 177
Dh 13106 CCGGCGGCGCGCTGCTGCTCTCGGAGACATCATCTCGGCTTGATCCCGGAGAGTAC 13047
Qy 178 GlyGlyMetProValSerGlyAspArgTTPProThrGlyLeuArgIleCysLeuAlaGlu 197
Dh 13046 GCGCGCGGT-----TGG-----ACGGGCGACGACCGCCCATACCTTGAAC 13008
Qy 198 GlnLeuLeuGlnSerLeuArgAlaAlaGlyPheGlnIleLeuAspTTPGluAspValSer 217
Dh 13007 AGCTTACGCGCGCTGTCAGCGAAGCCGGGTTCGAGATCTCGAAGTCACCGAGCTCAG 12948
Qy 218 SerArgThrArgTyrPheMetProGlnPheAlaGluGluLeu----- 231
Dh 12947 GCACAGACCGAGTGCATGCTCTCTGTAGTCGAGAGTTGCTCGGAAACTCGATGAG 12888
Qy 232 -----AlaAlaHisGlnHis-----GlyIle 238
Dh 12887 CTCGCGCGCGCTGAGCTGCGGCTGTCGACCTACGACGACGATCTTGGAGAGCATC 12828
Qy 239 AlaAspArgTyrGlyProAlaValAlaGlyTTPAlaAlaValCysAspTyrGluLys 258
Dh 12827 GCGGCGAAGCAGCGAGCCGAGCAGCAGCTGATCGCGGTGGGAAATACCGGAAA 12768
Qy 259 -----TyrAlaHisAsp-----MetGlyTyrAlaIleLeuThrAlaArgLys 272
Dh 12767 CATCCGGATTACCGCAGAAACGAGAAAGATGGGTTTATGCTCTCGAGGCTCGAAAG 12708
```

```
RESULT 6
US-09-603-207-1/c
; Sequence 1, Application US/09603207B
; Patent No. 6521406
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/603,207B
; EARLIER FILING DATE: 2000-06-23
; EARLIER APPLICATION NUMBER: 09/370,700
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 80161
; TYPE: DNA
; ORGANISM: Saccharopolyspora spinosa
US-09-603-207-1
```

```
Alignment Scores: 1.86e-39 Length: 80161
Pred. No.: 435.50 Matches: 108
Score: 51.67% Conservative: 47
Percent Similarity: 36.00% Mismatches: 96
Best Local Similarity: 30.16% Indels: 49
Query Match: 4 Gaps: 11
DB: 11
```

US-10-069-353a-8 (1-275) x US-09-603-207-1 (1-80161)

```
Qy 7 ProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThrProLeuLeuAsnSerVal 26
Dh 13562 CCGACCGCGGATCAGGTGAGATTCATCTTCATGCGTTG-----GCG 13521
Qy 27 AlaGlyGlyProCysAlaAlaIleHisGlyTyrTP-----GluAsnAsp 41
Dh 13520 CAGGGCGGTC-----CTGACACCGGTTACTGGGCGGGGATGCGAGGATGCGC 13467
Qy 42 GlyArgAlaSerTTPGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArgThr 61
Dh 13466 GGTGCCACACCGTGGTGGATCTGCGACCAACTGACCGACCTGTTCATCAACAAGGCC 13407
Qy 62 ValLeuAspGlyGlyValArgLeuLeuAspValGlyGlyThrGlyGlnProAlaLeu 81
Dh 13406 GCGCTCCGTCGCGGAGCGGACGCTTTCACCTCGGCTCGGCAATGGGACGCGTATGTC 13347
Qy 82 ArgValAlaAspAspAlaIleGlnIleThrGlyIleThrValSerGlnValGlnVal 101
Dh 13346 CGTGGCGGATGCGCAGCGCGGCTTGATCACCAGATCACCGGATCACCGGACCGCATCTC 13287
Qy 102 AlaIleAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspHisSerGly 121
Dh 13286 GCGCGCGCCACGAGCTGCGCAGACGAGACCGGACCTGGCGGAGTCTTGATTCATCTTA 13227
Qy 122 ValAaPAlaMeSerLeuProTyrProAspAaAlaPheAspAlaAlaTTPAlaMeGln 141
Dh 13226 GTCGAGCGGCGCCGAGCTGCGCTTACCGGAGCTTTCTTTCAGGCGCGATGGCGATGACG 13167
Qy 142 SerLeuLeuGlnMeSerGluProAspAaAlaAlaArgGluIleLeuArgValLeuLys 161
Dh 13166 TCCGTGTGCAGATCTGGACGACGCGCGGATCGGAGGTCCAGAGTCCAGAAATCTCGGA 13107
Qy 162 ProGlyGly-----IleuGly-----ValThrGluValValLysArgGluAlaGly 177
Dh 13106 CCGGCGGCGCGCTGCTGCTCTCGGAGACATCATCTCGGCTTGATCCCGGAGAGTAC 13047
Qy 178 GlyGlyMetProValSerGlyAspArgTTPProThrGlyLeuArgIleCysLeuAlaGlu 197
Dh 13046 GCGCGCGGT-----TGG-----ACGGGCGACGACCGCCCATACCTTGAAC 13008
Qy 198 GlnLeuLeuGlnSerLeuArgAlaAlaGlyPheGlnIleLeuAspTTPGluAspValSer 217
Dh 13007 AGCTTACGCGCGCTGTCAGCGAAGCCGGGTTCGAGATCTCGAAGTCACCGAGCTCAG 12948
Qy 218 SerArgThrArgTyrPheMetProGlnPheAlaGluGluLeu----- 231
Dh 12947 GCACAGACCGAGTGCATGCTCTCTGTAGTCGAGAGTTGCTCGGAAACTCGATGAG 12888
Qy 232 -----AlaAlaHisGlnHis-----GlyIle 238
Dh 12887 CTCGCGCGCGCTGAGCTGCGGCTGTCGACCTACGACGACGATCTTGGAGAGCATC 12828
Qy 239 AlaAspArgTyrGlyProAlaValAlaGlyTTPAlaAlaValCysAspTyrGluLys 258
Dh 12827 GCGGCGAAGCAGCGAGCCGAGCAGCAGCTGATCGCGGTGGGAAATACCGGAAA 12768
Qy 259 -----TyrAlaHisAsp-----MetGlyTyrAlaIleLeuThrAlaArgLys 272
Dh 12767 CATCCGGATTACCGCAGAAACGAGAAAGATGGGTTTATGCTCTCGAGGCTCGAAAG 12708
```

```
RESULT 7
US-09-029-603-4
; Sequence 4, Application US/09029603
; Patent No. 6210935
; GENERAL INFORMATION:
; APPLICANT: Schnupp, Thomas
; APPLICANT: Engel, Natalie
; APPLICANT: Bietenhader, Jurg
; APPLICANT: Toupet, Christine
; APPLICANT: Posdziech, Andreas
; TITLE OF INVENTION: Staurosporin Biosynthesis Gene Clusters
; FILE REFERENCE: 4-20555/A/PT
```



```

QY      45 SerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGlyArgThrValLeuAsp 64
      985 ACCATCGAGAGCGCGTGAACCGAGCTACCGAGTGTTCATGAGAACGCGTGAACGCGTGC 1044
QY      65 GtGlyValArgLeuLeuAspValGlyCyGlyThrGlnProAlaLeuArgValAla 84
      1045 GCCACCTCCACAGTCTCTGACCTCGGCTCGGGGTGGGGGGCCCGGCTCAAGGGTCTGG 1104
QY      85 ArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGlnValAlaIleAla 104
      1105 GCGCGCACCGGGGACGGGTCAACCGGCATCAGCATGACGAGGACGAGGACGAGCGCCG 1164
QY      105 AlaAspCysAlaArgGlnArgGlyLeuSerHisArgValAspPheSerCysValAspAla 124
      1165 AACCGGCTGGCGCGCGAGCGGGGTGCGCGACCGGTGCGGTTCAGCATGGCGACGCG 1224
QY      125 MetSerLeuProTyProAspAsnAlaPheAspAlaAlaTrpAlaMetGlnSerLeuLeu 144
      1225 ATGAACCTGCTCCCTCCGACGCGCTGTCGACGCGGTGATGGCGGTGATGATCTGC 1284
QY      145 GlnMetSerGlnProAspArgAlaIleArgGlnIleLeuArgValLeuLysProGlyGly 164
      1285 CACATGCGCGACCGGACGAGGTGTTCACCGAGGTGTGCGGCTGCTGCGCGCGCGGCG 1344
QY      165 IleLeuGlyValThrGlnValValLysArgGlnAlaGlyGlyLysMetProValSerGly 184
      1345 CGGATGCTCTCACCGACATCTTCGAGCGCCAC-----CGCGCAAGGCG 1389
QY      185 AspArgTrpProThrGlnLeuArgGlyLeu-----Ala 196
      1390 GTTCGACACCCCGGCGATCGACAAAGTTCGCGGACGTGATGACACCGGCGAATC 1449
QY      197 GlnGlnLeuLeuGlnSerLeuArgAlaAlaGlyPheGlnIleLeuAspTrpGlnAspVal 216
      1450 GACGACTACGTGGCGCTGTCGACCGCTCGGCTGGCGGCGGAGATCGTGCACGCTG 1509
QY      217 SerSerArgThrArgGlyTrpPheMetProGlnPheAlaGlnLeu-----Ala 232
      1510 ACCGACGACGACCGCTG-----CGCTCGCGACGAGATCGGACGCTCGCGGCG 1560
QY      233 AlaHisGlnIleGlyIleAla-----AspArgTyr 242
      1561 GTCGAGAGCGCGCGCGTGGCCATGAGACGAGGCAATTCGCTCGGCGACGATCTTTC 1620
QY      243 GlyProAla---ValAlaGlyTrpAlaAlaAlaValCysAspTyrGlnLysTyrAlaHis 261
      1621 AAGCCGTCGACGCTCGGCGGCG-----GTGAC 1647
QY      262 AspMetGlyTyrAlaIleLeuThrAlaArgLysPro 273
      1648 GACTTCGGCTGCTCTCTGTCACGCGCGACGCGCGCC 1683
      DB
      RESULT 12
      US-09-885-723-21
      ; Sequence 21, Application US/09885723
      ; Patent No. 6822142
      ; GENERAL INFORMATION:
      ; APPLICANT: Monsanto Company
      ; TITLE OF INVENTION: TRANSGENIC PLANTS CONTAINING ALTERED LEVELS OF STEROID COMPOUNDS
      ; FILE REFERENCE: WPC6783.1
      ; CURRENT APPLICATION NUMBER: US/09/885.723
      ; NUMBER OF SEQ ID NOS: 33
      ; SOFTWARE: PatentIn version 3.0
      ; SEQ ID NO 21
      ; LENGTH: 1444
      ; TYPE: DNA
      ; ORGANISM: Nicotiana tabacum
      ; US-09-885-723-21
      Alignment Scores:
      Pred. NO.: 7.29e-19 Length: 1444
      Score: 241.50 Matches: 71

```

```

Percent Similarity: 49.15% Conservative: 45
Best Local Similarity: 30.08% Mismatches: 88
Query Match: 16.72% Indels: 35
DB: 4 Gaps: 4
US-10-069-353a-8 (1-275) x US-09-885-723-21 (1-1444)
QY      9 SerGlnGlnValGlyGlnMetTyrAspLeuValThrProLeuLeuAsnSerValAlaGly 28
      253 ACTGACATGTTTAACAAATACATATCTTTGACATAG-CTTCTACG-AAATCGCGCTGGG 310
QY      29 GlyProCysAlaIleHisIleHisGlyTyrTrpGlnAsnAspGlyArgAlaSerTrpGlnGln 48
      311 GAGAGTCATTCATTTTGACACC-AGTGGAAAGAGAA-----TCACTCCAAAGG 360
QY      49 AlaAlaAspArgLeuThrAspLeuValAlaGlnArgThrValLeuAspGlyGlyValArg 68
      361 AGCATTTAAAGGCATGACGACCTTTCTGCTTGGCACTGGGATTTGAAACGAGCAAAAG 420
QY      69 LeuLeuAspValAlaGlyCyGlyThrGlyGlnProAlaLeuArgValAlaArgAspAsnAla 88
      421 GTCTTGACGTGAGATGATGGAATTTGGTGGCGCTTAAGAGAAATTCCTGATTCAGCTCT 480
QY      89 IleGlnIleThrGlyIleThrValSerGlnValGlnValAlaIleAlaAlaAspCysAla 108
      481 ACATGATGTACAGGCGCTCAACAATATGATATCAGATCTTCAAGGCGACAGGTGTTCAC 540
QY      109 ArgGlnValGlyLeuSerHisArgValAspPheSerCysValAspAlaMetSerLeuPro 128
      541 CCCAAAGTGGATGGATGATGACGACTTTCGCTTGTAAAGGATTTCAATGAAAGGCCA 600
QY      129 TyrProAspAsnAlaPheAspAlaAlaTrpAlaMetGlnSerLeuLeuGlnMetSerGln 148
      601 TTCCTGACAAATAGCTTTGATGACAGTATGCGAATATGAGAACTCACTGCATGACCAAT 660
QY      149 ProAspArgAlaIleArgGlnIleLeuArgValLeuLysProGlyGlyIleLeuGlyVal 168
      661 CCATTGGATGTATTAAGATTTTACCGGGGTGTGAACCTGTGATGATGTTTCGCTGTG 720
QY      169 ThrGln-----Val 171
      721 TATGAGTGTGTCATGACCGATCTTTCACCCCAATATGAAAGACACACAGATCAAG 780
QY      172 ValLysArgGlnAlaGlyGlyMetProValSerGlyAspArgTrpProThrGlyLeu 191
      781 GCCGAATTTGAGCTCGAAATGGCTTCTCT----- 810
QY      192 ArgIleCysLeuAlaGlnGlnLeuGlnSerLeuArgAlaAlaGlyPheGlnIleLeu 211
      811 GAGGTTAGATTGACACACAGATGCTCGAAGCAGCCAAACAGCTGTTTGAAGTTGTA 870
QY      212 AspTrpGlnAspValSer-----SerArgThrArgTyrPheMetPro 225
      871 TGGACAAGATCTGGCTGATGATCACTGTTCCATGTGTAATTGGCT 918
      DB
      RESULT 13
      US-09-248-796A-3873
      ; Sequence 3873, Application US/09248796A
      ; Patent No. 6747137
      ; GENERAL INFORMATION:
      ; APPLICANT: Keith Weinstock et al
      ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
      ; FILE REFERENCE: 107196.132
      ; CURRENT APPLICATION NUMBER: US/09/248.796A
      ; PRIOR FILING DATE: 1999-02-12
      ; PRIOR APPLICATION NUMBER: US 60/074.725
      ; PRIOR FILING DATE: 1998-02-13
      ; PRIOR APPLICATION NUMBER: US 60/096.409
      ; PRIOR FILING DATE: 1998-08-13
      ; NUMBER OF SEQ ID NOS: 28208
      ; SEQ ID NO 3873
      ; LENGTH: 1164

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 29, 2005, 19:51:25 ; Search time 699 Seconds

(without alignments)
2574.177 Million cell updates/sec

Title: US-10-069-353a-8

Perfect score: 1444

Sequence: 1 VLEGGAPTSQGVQGMVLDVT.....YERYADMGVAILTARKPVG 275

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=x1h
-O/cgn2_1/USPTO.apool/US10069353.rnact.26082005.172055.14723/app.query.fasta_1.455
-DB=Published Applications NA -QMT=fastap -SUPFIX=rnpb -MINMATCH=0.1
-LOPCU=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62
-TRANS=human40.cci -DIST=45 -DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAKLEN=200000000 -USER=US10069353 @CGN 1 1 480 @rnact.26082005.172055.14723
-NCPU=6 -ICPU=3 -NO MAP -LARGECURRY -NEG SCORES=0 -WMIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/PCr_NEW_PUB.seq:*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
6: /cgn2_6/prodata/1/pubpna/PCrUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/1/pubpna/US09C_NEW_PUB.seq:*
13: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
19: /cgn2_6/prodata/1/pubpna/US10F_PUBCOMB.seq:*
20: /cgn2_6/prodata/1/pubpna/US10F_PUBCOMB.seq:*
21: /cgn2_6/prodata/1/pubpna/US10H_PUBCOMB.seq:*
22: /cgn2_6/prodata/1/pubpna/US10H_PUBCOMB.seq:*
23: /cgn2_6/prodata/1/pubpna/US10I_PUBCOMB.seq:*
24: /cgn2_6/prodata/1/pubpna/US10I_PUBCOMB.seq:*
25: /cgn2_6/prodata/1/pubpna/US10J_PUBCOMB.seq:*
26: /cgn2_6/prodata/1/pubpna/US10J_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1444	100.0	80161	17	US-10-329-148A-1	Sequence 1, Appl
2	441.5	30.6	849	15	US-10-156-761-925	Sequence 925, Appl
3	441.5	30.6	9025608	15	US-10-156-761-1	Sequence 1, Appl
4	438.5	30.4	846	16	US-10-132-134-21	Sequence 21, Appl
5	438.5	30.4	52101	16	US-10-132-134-1	Sequence 1, Appl
6	435.5	28.2	80161	17	US-10-329-148A-1	Sequence 1, Appl
7	412	28.5	88400	21	US-10-844-716-1	Sequence 36, Appl
8	376.5	26.1	852	10	US-09-953-348-36	Sequence 36, Appl
9	376.5	26.1	852	15	US-10-267-255-36	Sequence 36, Appl
10	376.5	26.1	53500	10	US-09-953-348-76	Sequence 76, Appl
11	376.5	26.1	53500	15	US-10-267-255-76	Sequence 76, Appl
12	361.5	25.0	30000	11	US-09-980-217-1	Sequence 35, Appl
13	345	23.9	828	10	US-09-953-348-35	Sequence 35, Appl
14	345	23.9	828	15	US-10-267-255-35	Sequence 35, Appl
15	268	18.6	52659	21	US-10-496-377-1	Sequence 1, Appl
16	250	17.3	843	16	US-10-279-029-54	Sequence 54, Appl
17	250	17.3	843	16	US-10-219-810-17	Sequence 17, Appl
18	250	17.3	858	17	US-10-369-493-42625	Sequence 42625, A
19	246.5	17.1	843	16	US-10-279-029-53	Sequence 53, Appl
20	246.5	17.1	843	16	US-10-219-810-16	Sequence 16, Appl
21	242.5	16.8	1041	19	US-10-257-344A-13	Sequence 13, Appl
22	241.5	16.7	1444	10	US-09-885-723-21	Sequence 21, Appl
23	241.5	16.7	1444	21	US-10-862-907-21	Sequence 21, Appl
24	240	16.6	930	16	US-10-219-810-44	Sequence 44, Appl
25	238.5	16.5	1386	18	US-10-425-114-20190	Sequence 20190, A
26	238.5	16.5	1582	18	US-10-425-114-3622	Sequence 3622, Ap
27	238.5	16.5	1595	18	US-10-425-114-29071	Sequence 29071, A
28	238	16.5	1370	18	US-10-425-114-14934	Sequence 14934, A
29	238	16.4	1551	18	US-10-425-114-14931	Sequence 14931, A
30	237	16.4	1457	9	US-09-779-144A-6	Sequence 6, Appl
31	236.5	16.4	1457	9	US-10-437-963-65029	Sequence 65029, A
32	234.5	16.2	1230	19	US-10-437-963-45285	Sequence 45285, A
33	232.5	16.1	1089	19	US-10-437-963-45285	Sequence 45285, A
34	232.5	16.1	1630	19	US-10-437-963-45286	Sequence 45286, A
35	232	16.1	1044	16	US-10-219-810-8	Sequence 8, Appl
36	232	16.1	2590	17	US-10-418-552-151	Sequence 151, App
37	229.5	15.9	681	21	US-10-487-901-2303	Sequence 6303, Ap
38	229.5	15.9	681	21	US-10-487-901-6303	Sequence 43, Appl
39	229.5	15.9	1038	16	US-10-279-029-43	Sequence 4, Appl
40	229.5	15.9	1038	16	US-10-219-810-4	Sequence 1, Appl
41	229.5	15.9	1038	19	US-10-668-240-6	Sequence 6, Appl
42	229.5	15.9	7713	19	US-10-668-240-5	Sequence 5, Appl
43	229.5	15.9	8149	19	US-10-668-240-5	Sequence 9, Appl
44	228	15.8	1074	16	US-10-219-810-9	Sequence 44, Appl
45	226	15.7	1131	16	US-10-279-029-44	Sequence 44, Appl

ALIGNMENTS

RESULT 1
US-10-329-148A-1
; Sequence 1, Application US/10329148A
; Publication No. US20040023343A1
GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 Div1
; CURRENT APPLICATION NUMBER: US/10/329,148A
; CURRENT FILING DATE: 2002-12-23

```
/ PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/603,207B
/ PRIOR FILING DATE: EARLIER FILING DATE: 2000-06-23
/ PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/370,700
/ PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-09
/ NUMBER OF SEQ ID NOS: 39
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 1
/ LENGTH: 80161
/ TYPE: DNA
/ ORGANISM: Saccharopolyspora spinosa
US-10-329-148A-1

Alignment Scores:
Pred. No.: 1,29e-160 Length: 80161
Score: 1444.00 Matches: 275
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-10-069-353A-8 (1-275) x US-10-329-148A-1 (1-80161)

QY 1 ValLeuProGlyValAlaProThSerGlnGlnValGlyGlnMetTyraSpleuValThr 20
DB 20168 GTGTTGCCAGTGGCCGACCAACATCGCAGAGTTGGGCAAGTATGACCTGGTCAAG 20227
QY 21 ProLeuLeuSerValAlaGlyProCySaAlaIleHisGlyTyrTrpGluAsn 40
DB 20228 CCGTGTCTAACTCGGTCCGGGGCGCCCTGTGGCCATCCACACGCGTACGAGGAAC 20287
QY 41 AspGlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
DB 20288 GACGGCGCGGCTTCCTGGAGAGAGCGCGCAGCCGCTACACGACTTGTCCCGAAGC 20347
QY 61 ThrValLeuAspGlyValAlaArgLeuLeuAspValGlyCyGlyThrGlyGlnProAla 80
DB 20348 ACCGGTCTGATGGCGCGCTGACCTGATGTGGGGTGGCGTACCGACACCGAGCG 20407
QY 81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGln 100
DB 20408 CTGGCGGTGGCGCGACACCGGATCCGATCACCGGATCACCGTCCGCCAGGTGCA 20467
QY 101 ValAlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120
DB 20468 GTGGCCATGCGCGCTGATGGCGCAGACCGGACCTAAGCCACCGGGTGGACTTCTG 20527
QY 121 CysValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
DB 20528 TACGTCGATGCGATGTCCCTGCGGTAACCGGACAACTTTCGACGCGCTGGGCGCATG 20587
QY 141 GlnSerLeuLeuGlnMetSerGluProAspArgAlaIleArgGluIleLeuArgValLeu 160
DB 20588 CAGTGCCTTTGAGATGTCCCAACCGACCGGTCATCCGGGAATCTTTCGAGTACTC 20647
QY 161 LysProGlyGlyIleLeuGlyValThrGluValIleLysArgGluAlaGlyGlyMet 180
DB 20648 AAACCCGGGTGATCTCGCGCTCACCGAGTGTCTAAACGAGAAAGCGGCGCGGAGT 20707
QY 181 ProValSerGlyAspArgTrpProThrGlyLeuAlaGlyLeuAlaGlnLeuLeu 200
DB 20708 CCGGTGTCCGGGACAGGTGGCGACCGGCTTCGATGTGCTCGTGAAGCACTTCTG 20767
QY 201 GlnSerLeuArgAlaAlaGlyPheGluIleLeuAspTrpGluAspValSerSerArgThr 220
DB 20768 GAATGCTGTGTCGACGCGGGGTTCCAGATCTCTCGATTGGAGAGAGCTGTCTCGAGAAC 20827
QY 221 ArgTyrPheMetProGlnPheAlaGlnGluLeuAlaAlaHisGlnHisGlyIleAlaAsp 240
DB 20828 CCGTACTTCATGCGCGAGTTCGCCGAAAGCTCGCTGGCACACAGACGAGGATCCGGAC 20887
QY 241 ArgTyrGlyProAlaValAlaGlyTrpAlaAlaAlaValCysAspTyrGlnLysTyrAla 260
DB 20888 AGTACGGGCGCGCTGTCTGCGCGGCTGGGCGCGGCTGTGCAATTATGAGAAATATGCG 20947
```

```
QY 261 HisAspMetGlyTyrAlaIleLeuThrAlaArgLysProValGly 275
DB 20948 CACGACATGGCTATGCAATTCATGACGGCGGAAAGCCGGTGGC 20992

RESULT 2
US-10-156-761-925
/ Sequence 925, Application US/10156761
/ Publication No. US20030119018A1
/ GENERAL INFORMATION:
/ APPLICANT: OMURA, SATOSHI
/ APPLICANT: IKEDA, HARUO
/ APPLICANT: ISHIKAWA, JUN
/ APPLICANT: HORIKAWA, HIROSHI
/ APPLICANT: SHIBA, TADAYOSHI
/ APPLICANT: SAKAKI, YOSHIYUKI
/ APPLICANT: HATTORI, MASAHIRA
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-262
/ CURRENT APPLICATION NUMBER: US/10/156,761
/ PRIOR FILING DATE: 2002-05-29
/ PRIOR APPLICATION NUMBER: JP 2001-204089
/ PRIOR FILING DATE: 2001-05-30
/ PRIOR APPLICATION NUMBER: JP 2001-272697
/ PRIOR FILING DATE: 2001-08-02
/ NUMBER OF SEQ ID NOS: 15109
/ SEQ ID NO 925
/ LENGTH: 849
/ TYPE: DNA
/ ORGANISM: Streptomyces avermitilis
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(849)
US-10-156-761-925

Alignment Scores:
Pred. No.: 7,65e-43 Length: 849
Score: 441.50 Matches: 111
Percent Similarity: 56.32% Conservative: 45
Best Local Similarity: 40.07% Mismatches: 107
Query Match: 30.57% Indels: 15
DB: 15 Gaps: 6

US-10-069-353A-8 (1-275) x US-10-156-761-925 (1-849)

QY 7 ProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThrProLeuLeuSerVal 26
DB 28 CCTCTTCCCTGGAAGTGGGGACTACTACGACCGTTTGACCGACCTCATGAAATCGTGGC 87
QY 27 AlaGlyGlyProCysAlaIleHisGlyTyrTrpGluAsnAspGlyArgAlaSerTrp 46
DB 88 CTGGGTGG-----MACACCACTCGGATATCTGGCGCGACCGGGGAGCGGACATTCA 141
QY 47 Gln---GlnAlaAlaAspArgLeuThrAspLeuValAlaGluArgThrValLeuAspGly 65
DB 142 CCGGGCAAGCGCGCGACCGGCTCACCGATCTCTCATCGGCAAGCTGAGAGGATCAGC 201
QY 66 GlyValAlaGluLeuLeuAspValGlyCyGlyThrGlyGlnProAlaLeuArgValAlaArg 85
DB 202 GCGCGCGGCTGAGAGTGGGCTGGGCTCCGAAAGCGGAGCGGCGGCTGGCCCTG 261
QY 86 AspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGlnValAlaIleAla 105
DB 262 AGCGCGCCCTCGATGTCTGGGCTGACGTTGACGATTCAGTTCGGGCTGGCCAGC 321
QY 106 AspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSerCysValAspAlaMet 125
DB 322 GCTCTCGGAAGACGTGCGACGTGGCGGACCGGGTGTGTTTCCACCGGCGGACGCGATG 381
QY 126 SerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMetGlnSerLeuLeuGlu 145
DB 382 GAATCGCGTTTCCCGACGGGTCTTTCGACCGGCGGTGGGCGCGGAGTGTCTCTGAC 441
```

```
QY 146 MetSerGluProAspArgAlaIleArgGluIleLeuArgValLeuLysProGlyGlyIle 165
DB 442 ATGCCAGCCCGCCGACAGGATGATCGGAGATCGCCGGGTCTCCGCCCCCGCCGCG 501
QY 166 LeuGlyValThrGluValValLysArg-GluAlaGlyGlyLysMetProValSerGlyLys 185
DB 502 CTGGCCGTCAGGACGTCGACCTGGCGGCTTGGGCGGACCGGACGTAAGCCGGGGAG 561
QY 185 pArgTrpProThiGlyLeuArgIle-----CysLeuAlaGluGlnLeuLysIle 202
DB 562 TGACACTCCCAAGTGTCTGGCGGTCGCGGCTGTGCACATGACAGATACGCGGAAATG 621
QY 202 rLeuArgAlaAlaGlyPheGluIleLeuSerTrpGluAspValSerSerArgThrArgTrp 222
DB 622 -ATCGCCGACGCGCGGCTGAACTGATGCTGACGACGACGACGACGACGACGACGACG 675
QY 222 rPheMetProGlnPheAla-----GluGluLeuAlaAlaIleGlnIleGlyIleAl 239
DB 676 -GTGGGCCCCCTCTTGTGCGCGCGCTGTCACAGTGAACAGACACTCGACAGATACGC 734
QY 239 aAspArgTrpGlyProAlaValAlaGlyTrp-----AlaAlaAlaValCysAspTrpG 257
DB 735 GGGCGGCTTCGGATGGCGGTCGCGGAGATGCGGAAGTGTGTTCACAGTGAACAGCG 794
QY 257 uLysTrpAlaIleAspMetGlyTrpAlaIleLeuThrAlaArgLysPro 273
DB 795 CCCCTGACGCGCGACATCGGCTATGTCTGATGACGCGCGCGCGCG 843

RESULT 3
US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156, 761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; NAME/KEY: misc_feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Alignment Scores:
Pred. No.: 1.16e-37 Length: 9025608
Score: 441.50 Matches: 111
Percent Similarity: 56.32% Conservative: 45
Best Local Similarity: 40.07% Mismatches: 107
Query Match: 30.57% Indels: 15
DB: 15 Gaps: 6

US-10-069-353a-8 (1-275) x US-10-156-761-1 (1-9025608)
QY 7 ProThrSerGlnGlnValGlyGlnMetTrpAspLeuValThrProLeuLeuAsnSerVal 26
DB 1137453 CCCCTCTCCCTGGAAGTGGGAGACTACTACGACCGTTTGAACGACCTCATGAACTCGCGC 1137394
```

```
QY 27 AlaGlyGlyProCysAlaIleIleHisIleGlyTrpTrpGluAsnAspGlyArgAlaSerTrp 46
DB 1137393 CTGGGTGGG-----AACACCACTCTGGATCTGAGCCGCGACCGGGGAGCGGATTC 1137340
QY 47 Gln---GlnAlaAlaAspArgLeuThrAspLeuValAlaGluArgThrValLeuAspGly 65
DB 1137339 CCGGGGAAGCGCGCGACCGGCTCAACCGATCTCTCATCGACAGCTGAGAGGATCAACG 1137280
QY 66 GlyValAspLeuLeuAspValGlyCysGlyThrGlyGlnProAlaLeuArgValAlaArg 85
DB 1137279 GCGCGCGGCTCTGGAACCTCGGCTGCGGCTTCCGGAAGCCGCGCGGCTGCGCTG 1137220
QY 86 AspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGlnValAlaIleAlaAla 105
DB 1137219 AGCGGCGCGCGTGAATGCTCGGCTGACGAGTGAAGGAGTTCAGGTCGGGCTGGAGAC 1137160
QY 106 AspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSerCysValAspAlaMet 125
DB 1137159 GCTCTCGGAGACAGTGCAGCGTGGCGGACCGGGTCTGTTCACCCGTCGACCGGATG 1137100
QY 126 SerLeuProTrpProAspAsnAlaPheAspAlaAlaTrpAlaMetGlnSerLeuGlu 145
DB 1137099 GAATGCTCCGTTCCCGACGGGTCCTTCAGACGGGCTGGGCTGGAGTCTCTGACAC 1137040
QY 146 MetSerGluProAspArgAlaIleArgGluIleLeuArgValLeuLysProGlyGlyIle 165
DB 1137039 ATGCCAGCCCGCCGACAGTGAATCGGAGATCGCCGGGTCTCCGCCCCCGCGCGCG 1136980
QY 166 LeuGlyValThrGluValValLysArg-GluAlaGlyGlyLysMetProValSerGlyLys 185
DB 1136979 CTGGCGGTCAGGACGTCGACATGCGCGCTTCCGGCGGACCGGACGTAAGCCGCGGAG 1136920
QY 185 pArgTrpProThiGlyLeuArgIle-----CysLeuAlaGluGlnLeuLysIle 202
DB 1136919 TGCACGTCCTCAGTGTCTCCGCGCGCGCTGTCACATGACAGATACCGCGGAAATG 1136860
QY 202 rLeuArgAlaAlaGlyPheGluIleLeuSerTrpGluAspValSerSerArgThrArgTrp 222
DB 1136859 -ATGCCGACGCGCGGCTGAACTGATGATGACGACCGACCATCGCGGATCAGGTC- 1136806
QY 222 rPheMetProGlnPheAla-----GluGluLeuAlaAlaIleGlnIleGlyIleAl 239
DB 1136805 -GTGGGCCCCCTCTTGTGCGCGCGCTGTCACAGTGAACAGACACTCGACAGATACGC 1136747
QY 239 aAspArgTrpGlyProAlaValAlaGlyTrp-----AlaAlaAlaValCysAspTrpG 257
DB 1136746 GCGCGGCTTCGGATGCGCGTCGCGGAGATGCGGAAGTGTGTTCACAGTGAACGCT 1136687
QY 257 uLysTrpAlaIleAspMetGlyTrpAlaIleLeuThrAlaArgLysPro 273
DB 1136686 CCCCTGACGCGCGACATCGGCTATGTCTGATGACGCGCGCGCGCG 1136638

RESULT 4
US-10-132-134-21
; Sequence 21, Application US/101322134
; Publication No. US20030171562A1
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Yang, Xianshu
; APPLICANT: Staffe, Alfredo
; APPLICANT: Zazopoulos, Emmanuel
; TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES
; FILE REFERENCE: 3012-205
; CURRENT APPLICATION NUMBER: US/10/132,134
; PRIOR FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 846
; TYPE: DNA
; ORGANISM: Streptomyces platensis subsp. roseaceus
US-10-132-134-21
```

Alignment Scores:

Pred. No.:	1.75e-42	Length:	84
Score:	438.50	Matches:	10
Percent Similarity:	56.10%	Conservative:	55
Best Local Similarity:	36.93%	Mismatches:	95
Query Match:	30.37%	Indels:	31
DB:	16	Gaps:	9

US-10-069-353A-8 (1-275) X US-10-132-134-21 (1-846)

QY	2	LeuProG1YGIYLaIaProThiSerInGInVa1GIYGIYmeTYrAspLeuValThPro	21
Db	34	GTCCCGCTCCGCCCCCGCCCGAAGAGGTGGACACTCTAGACCGGCTCACCGCA	93
QY	22	LeuLeuhsuSerVala1aGIYGIYProCySa1a1IeH1sG1YrTYrTrpGlu--Asn	40
Db	94	CTGGACACCCGAGCGCGCGCGG-----ACCTTCACCTCGGGTACTCGGAGGTGCAC	147
QY	41	AspG1YArG1aSerTrpGInGIna1a1aAspAryLeuThrAspLeuVala1aGIYArg	60
Db	148	GACACGACACCCCCCTCGTGAAGGGGCGGACCGAGCTTACCGACACATACGACCGC	207
QY	61	ThrValLeuhsprGIYGIYVala1ArgLeuLeuAspVala1GIYrG1YrThG1YrProa1a	80
Db	208	CTCGCGATCGACACGAGGAGACGGGGCTTCGACGTCCGGCTGGAGTGGCGCACCGGCC	267
QY	81	LeuArgVala1aArgAspAspA1a1eGIn1IeThrG1Y1IeThrValSerGInVa1GIn	100
Db	268	ATCGCGATCCCGCGCGACCGCGCCCATGTCAAGGCGATCGCATAGCAAGACAG	327
QY	101	Vala1a1Ie1a1a1aAspCySa1aArG1aArG1aArg1YLeuSerTh1aArgVala1aAspSer	120
Db	328	ATCGCCCGCGCACCGCCCTCGCGCGAGCGCGCGCTGAGAGACCGGTGGATTCCGG	387
QY	121	CyeVala1aAsp1aMetSerLeuProTYrProAspAspA1a1eAspA1a1a1rPa1aMet	140
Db	388	CACGCCGACCGCCATGAGATCGGCTTCCCGACGACTCTTCGACGCGCGCATCGCATC	447
QY	141	GInSerLeuend1uMetSerG1uProAspA1a1eArgG1u1IeLeuArgValLeu	166
Db	448	GAATCGATCTTCACATGCCCGACCGGACCGAGCGAGTCTCGCGAGATCCCGCGTACTG	507
QY	161	LyAsProG1YGIY1IeLeuG1YValThrG1uVala1ValYrAsrG1u1aGIYGIYmet	180
Db	508	CGCCCCGCGCGCGCTGGTCTTCACCGACTTCTTCGAGCGCGC-----	552
QY	181	ProValSerG1YAspArgTrpProThrG1Y-----LeuArg	192
Db	553	CCCGTCCCGCGAGAGACGCCCGCGGTGACCGGCTCTCCGCACTTCATCATGACG	612
QY	193	1IeCyeLeu1aGIuGInLeuLeuGInuSerLeuA1a1aGIYpHe-----Glu	209
Db	613	CTGGCCCGGCGCGAGAGACTACGTGCCCATCTCGCGCGACGAGGCTCGCTGTGAG	672
QY	210	1IeLeuAspTrpG1uAspValSerSerArgThrArgTYr--PheMetProGInPheA	228
Db	673	CTCTTCGACATCACCGAGGAGCGTGGCTCGACTTCGAGCAGATGAGCAGGCGCTCC	732
QY	229	G1uG1uLeu---Ala1aH1sGInH1sG1Y1eAlaAspAryrG1YrProa1aVala1a	247
Db	733	CAGAGATGACAGCGCTTTCACACGACGAGCAGAGAGAGAAAGTTCAAGCCCC--	783
QY	248	G1YrPa1a1a1aVala1CySa1aPryrG1uYrYr1a1aH1aAspMetG1YrYr1a1e	267
Db	784	-----GCCATGATCGACGTCGACGAATC-----GGTCTCGTTCG	822
QY	268	LeuThrAlaArgYrProVal	274
Db	823	CTGACCGCCCAAAAGCCCTC	843

```
RESULT 5
US-10-132-134-1
; Sequence 1, Application US/10132134
```

Publication No. US20030171562A1

GENERAL INFORMATION:

APPLICANT: Farnet, Chris

APPLICANT:	Staffa, Alfr
APPLICANT:	yang, Alansh
APPLICANT:	

APPLICANT: Zazopoulos, Emmanue

1. TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES

FILE REFERENCE: 3012-2US

CURRENT APPLICATION NUMBER: US/10/132,134

NUMBER OF SEC TO VOC: 43
; CURRENT FILING DATE: 2002-04-26

; NUMBER OF SEQ ID NOS: 43
SOFTWAPE: Patent In version 3.0

; SOFTWARE:
;
; SEO ID NO 1

DEQ ID NO 4
LENGTH: 52101

MEMORANDUM FOR THE DIRECTOR
SUBJECT: DNA
TYPE: DNA
;

US-10-132-134-1

Alignment Scores:

Pred. No.:	3.5e-40	Length:	52101
Score:	438.50	Matches:	106
Percent Similarity:	56.10%	Conservative:	55
Best Local Similarity:	36.93%	Mismatches:	95
Query Match:	30.37%	Indels:	31
DB:	16	Gaps:	9

US-10-069-353A-8 (1-275) X US-10-132-134-1 (1-52101)

QY	1	LeuProG1yG1yAlaProThrSerGlnValAG1yGlnMetTyrAspLeuValThrPro	21
Db	51267	GTCCCGCTCTCCGCCCCCGTCCCGAAGAGTGTGGACACTTACAGACCGCTCCACCGCA	51326
QY	22	LeuLeuAsnSerValAlaAG1yG1yProCysValAlaMetAsnAG1yTyrTrrGlu--Asn	40
Db	51327	CTGCACACCGAAGCGCGCGCGGC-----AGCTTCCACTCGGCTACTGGGACGTGCAC	51380
QY	41	AspGlyTyrAlaSerTrrGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg	60
Db	51381	GACACAGACACCCCGCTCGTGAAAGAGGCGACCGGCTCACAGACAGATGACGACCGC	51440
QY	61	ThrValLeuAspG1yG1yValArgLeuLeuAspValAG1yCysG1yThrG1yAlaProAla	80
Db	51441	CTCGCGATGCACAGGAGCAGAGGGGTCTCGAGTCTGGCTGGAGATGGCGCAGCGCGC	51500
QY	81	LeuArgValAlaArgAspAsnAlaIleGlnIleThrG1yTyrThrValSerGlnValGln	100
Db	51501	ATCCGATGCGCCGGCGACCGCGCGCCATGTCAAGGCGATGCGCATGACAGACGACAG	51560
QY	101	ValAlaIleAlaAlaAspCysAlaArgGluArgG1yLeuSerAspArgValaAspPheSer	120
Db	51561	ATGGCCCGGCGACCGCCCTCGCCAGAGGGCGCGGCTGAGGAGACCGGTGGAGTTCCGG	51620
QY	121	CysValaAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTrrAlaMet	140
Db	51621	CAGCGCGAGCGCATGAGACTGGCTTCCCGACGACTCTTGAGAGCGCGCATCCGCAATC	51680
QY	141	GlnSerLeuLeuGlnMetSerGlnProAspArgAlaIleArgGlnIleLeuArgValLeu	160
Db	51681	GAGTCATCTTCCACATGCCCGACCGCGGAGCGGGTCTCGCCGAGATCGCGCGCTACTCG	51740
QY	161	IysProG1yG1yIleLeuG1yValThrGluValValIyIyArgGluAlaG1yG1yMet	180
Db	51741	CGCCCGGGCGCGCTGTGCTCCACACGACTTTCGAGCGCGGC-----	51785
QY	181	ProValSerG1yAspArgTrrProThrGly-----LeuArg	192
Db	51786	CCCGTCCCGCAGAGACGAGCCCGCGGTGACCGGCTCTCCGACTTCATCATGTACG	51845
QY	193	IleCysLeuAlaGluGlnLeuLeuGlnSerLeuAspAlaAlaG1yPhe-----Glu	209
Db	51846	CTGGCCCGGCGCGAGGACTACGTGCCATCTGCGCGAGAGCGAGGCGCTTGCTGTGAG	51905
QY	210	IleLeuAspTrrGluAspValSerSerArgThrArgTyr--PheMetProGlnPheAla	228

```

Db      51906 CTCCTGACATCCACGAGAGCGTGCCTGAGACCTTGAGCATGAGCCAGGCTCC 51965
      229 GtGluLeu-----AlaAlaHISGlnISGlyLeaIAspArgTyrGlyProAlaValAla 247
      51966 CAGGAGATCCAGACCGCTTCCAGACGAGGAGAGAGAAAGTTCAAGCCCC----- 52016
Qy      248 GtYTPAlaAlaAlaValCyAspTyrGtGlyGlyTyrAlaHISAspMetGlyTyrAlaIle 267
      52017 -----GCCCTCCATGATCGACGTCCAGCAATTC-----GGCTCCGTTCTG 52055
Qy      268 LeuThrAlaArgIAspProVal 274
      52056 CTGACCGCCCAAAAGCCCTC 52076

RESULT 6
US-10-329-148A-1/c
; Sequence 1, Application US/10329148A
; Publication No. US20040023343A1
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madhuri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 Div1
; CURRENT APPLICATION NUMBER: US/10/329,148A
; PRIOR FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/603,207B
; PRIOR FILING DATE: EARLIER FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/370,700
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 80161
; TYPE: DNA
; ORGANISM: Saccharopolyspora spinosa
US-10-329-148A-1

Alignment Scores:
Pred. No.: 1,4e-39 Length: 80161
Score: 435.50 Matches: 108
Percent Similarity: 51.67% Conservative: 47
Best Local Similarity: 36.00% Mismatches: 96
Query Match: 30.16% Indels: 49
DB: Gaps: 11

US-10-069-353A-8 (1-275) x US-10-329-148A-1 (1-80161)
Qy      7 ProThSerGlnGlnValGlyGlnMetTyrAspLeuValThrProLeuLeuAAsnSerVal 26
      13562 CCGACCCCGCATCGAGTGGAGTCCATCTTCATCCGTTG-----GCG 13521
Db      27 AlaGlyGlyProCysAlaIleHISGlyTyrTrp-----GluAAsnAsp 41
      13520 CACGGGGCTGCC-----CTGCACCACTGTTACTGGCGGGCGGGGTATCCGAGAGATGCC 13467
Qy      42 GlyArgIAspTyrGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArgThr 61
      13466 GGTGCCACACCGTGTGGATGTCGCCACCACTGACCACTGTTCAATGACCAAGGCC 13407
Db      62 ValLeuAspGlyGlyValAlaGlyLeuAspValGlyCysGlyThrGlyGlnProAlaLeu 81
      13406 GCGCTCGCTCCCGGCGGACCTGTGACCTGGGCTGGCGCAATGGAGCCCGGTATGTC 13347
Qy      82 ArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGlnVal 101
      13346 GTGCGGCAATGCCGACGCGGCTTCGATGCCGGAATCACTCGAAGCCGACATCTC 13287

```

```

Qy      102 AlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHISArgValAspPheSerCys 121
      13286 GCCCGCCGACCAAGCTGCCAAGACCGGACTGGCCGGAGCTTGAGTTGATGTA 13227
Qy      122 ValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaIleTyrAlaMetGln 141
      13226 GTCCAGCGGCCCACTGCGCCCTTACCAGCGGTTCTTTCAGCGCGCATGGCGCATGACG 13167
Qy      142 SerLeuLeuGluMetSerGluProAspArgAlaIleArgGlyIleLeuArgValLeuArg 161
      13166 TCCGTCGTGCAGATGTCGACCAAGCGCCCGGATCCGAGATCCGAGATCCAGAACTCTGGA 13107
Qy      162 ProGlyGly-----IleLeuGly-----ValThrGluValAlaValAspArgGluAlaGly 177
      13106 CCGGCGCGCGGTTGCTGCTCCGAGACATCATCATCTCGGTTCCGATCCCGGAGATGAC 13047
Qy      178 GlyGlyMetProValSerGlyAspArgTyrProThrGlyLeuArgIleCysLeuAlaGlu 197
      13046 GCGGCGGCT-----TGG-----ACGGGACGACCGCCCATCTTGAAC 13008
Db      198 GlnLeuLeuGluSerLeuArgAlaAlaGlyPheGlyIleLeuAspTyrGluAspValSer 217
      13007 AGCTTCAACGCGCGCTGGTCAGCGAAGCCGGTTGAGATTCTGAAGTCAACGACCTCAG 12948
Qy      218 SerArgThrArgTyrPheMetProGlnPheAlaGluGluLeu----- 231
      12947 GCACAGACCAAGTGCATGTCTTCCTGTCGTCCGACGAGTTCTCCGGAACCTCATGAG 12888
Qy      232 -----AlaAlaHISGlnHIS-----GlyIle 238
      12887 CTCGCGCGCGGTGAGCGCTGCGGCTGTCCGACCTTACACCAAGCTTACGAGACATC 12828
Db      239 AlaAspArgTyrGlyProAlaValAlaGlyTyrAlaAlaAlaValCysAspTyrGlyLeu 258
      12827 GCGGCGAAGCAGCGACCGGACGACGACGATGTCGCGGTTGCGGAATGCCGGA 12768
Qy      259 -----TyrAlaHISAsp-----MetGlyTyrAlaIleLeuThrAlaArgGly 272
      12767 CATCCGATTCAGCCAGAAACAGAGAAAGCATGGTTTCAATGCTCTCGAGGCTCGAAG 12708

RESULT 7
US-10-844-716-1/c
; Sequence 1, Application US/10844716
; Publication No. US20050003409A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Chenglin
; APPLICANT: Chaleff, Deborah T.
; APPLICANT: Ruppen, Mark B.
; APPLICANT: Stephens, Jerome
; TITLE OF INVENTION: Cloning Genes From Streptomyces Cyaneogriseus Subsp.
; FILE REFERENCE: AM100484
; CURRENT APPLICATION NUMBER: US/10/844,716
; CURRENT FILING DATE: 2004-05-13
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 88400
; TYPE: DNA
; ORGANISM: bacteria
US-10-844-716-1

Alignment Scores:
Pred. No.: 1.05e-36 Length: 88400
Score: 412.00 Matches: 103
Percent Similarity: 50.88% Conservative: 42
Best Local Similarity: 36.14% Mismatches: 106
Query Match: 28.53% Indels: 34
DB: Gaps: 9

US-10-069-353A-8 (1-275) x US-10-844-716-1 (1-88400)
Qy      5 GlyAlaProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThrProLeuLeuAAsn 24

```



```
Db 51282 GGATATACCCACCCCGCGGACGCTGGTGACTACTACGACCGCATGACCGCTGTTGAAAC 51223
QY 25 SerValAlaGlyIYProCysAlaIleHisGlyTYTrpGluAsn---AspGlyArg 43
Db 51222 CGGGCCCTGGGGGGC-----AACAGCGACCTCGGATACCTGGCCGACCGGACGAGCGC 51169
QY 44 AlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArgThrValLeu 63
Db 51168 AACACACTCGGCGAGGCTCTCGACCGCTCACCGACCATATATGTCGGAAGCTTCGGGAA 51109
QY 64 AspGlyGly-----ValArgLeuLeuAspValGlyCysGlyTYThrGlyInProAla 80
Db 51108 CACACCGGGCGCCCGGTACAGAGTGTGTGACGTGCGCTGCGGTTCGCGCGCGCGCGC 51049
QY 81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGln 100
Db 51048 CTGGCTCTGGCCACAGCGACCGCGGTGACATGTCGCGCATCACCATCAGCCCGCGGACG 50989
QY 101 ValAlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120
Db 50988 GTCGAGCTGGCCACGCGCTCGCCGAGCGGTCCGGACTCGGAAACCGGGTCCGCTTCAG 50929
QY 121 CysValAspAlaMetSerLeuProTYTrpProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
Db 50928 TCGCGCGACGCGCATGAGCTGCGCTTCCCGACGCGCTTCCTCGACGCGCTCGGGCGCTTC 50869
QY 141 GlnSerLeuGluGlnMetSerGluProAspArgAlaIleArgGluIleLeuArgValLeu 160
Db 50868 GAGTCTCTCTCTCACAATGCCCGACCGCGCGCGGTTCACAGAGATGCGCCGTGTCTG 50809
QY 161 LysProGlyGlyIleLeuGlyValThrGluValValLysArgGluAlaGlyIleGlyMet 180
Db 50808 CCCCCCGGTGGCGGTGGCGCGCATGACGTCACTCCGCGCATCA----- 50761
QY 181 ProValSerGlyAspArgTrpProThrGlyLeuAlaGlyIleCysLeuAlaGlnLeu 200
Db 50760 -----CAGCCGACGGGGCGGAGCTGGTCTGCCAGCACTCCGCGCTC 50719
QY 201 GlnSerLeu-----ArgAlaAlaGlyPheGluIle 210
Db 50718 CCCTCGCTATCCCATCATCACCCTATGCGCGGAGATGATGACGAGCGCGGACTCCGGCTG 50659
QY 211 LeuAspTrpGluAspValSerSerArgThrArgTYRPhemeProGlnPheAla----- 228
Db 50658 ACCGAGCTCACCGCATCGGTGAGCAGCTC-----ATCGCCCGTCGTAACGCGCATG 50605
QY 229 ---GluGluLeuAlaAlaHisGlnIleGlyIleAlaAspArgTYRAlaValAla 247
Db 50604 GGTGACGAGTACGGCGCAACGCCCAAGCTTACGCGGAGCCCTGAGATGACGCGGAC 50545
QY 248 GlyTrpAlaAlaAlaVal-----CysAspTYRGlulysTYRAlaHisAspMetGlyTYR 265
Db 50544 GACCTGGAGACCTTCTGTGGGCAAGTGCAGC---CACTGTATACAGAGGACATCGGGTAC 50488
QY 266 AlaIleLeuThrAla 270
Db 50487 GTCTCTGTACCGCG 50473

RESULT 8
US-09-953-348-36
; Sequence 36, Application US/09953348
; Publication No. US20030134398A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, David. H
; APPLICANT: Mao, Yingqing
; APPLICANT: Varoglou, Mustafa
; APPLICANT: He, Min
; APPLICANT: Sheldon, Paul
; TITLE OF INVENTION: MITOMYCIN BIOSYNTHETIC GENE CLUSTER
; FILE REFERENCE: 600.530US1
; CURRENT APPLICATION NUMBER: US/09/953,348
; CURRENT FILING DATE: 2001-09-12
```

```
; PRIOR APPLICATION NUMBER: PCT/US00/06394
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/266965
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Streptomyces lavendulae
US-09-953-348-36

Alignment Scores:
Pred. No.: 4,86e-35 Length: 852
Score: 376.50 Matches: 101
Percent Similarity: 50.84% Conservative: 51
Best Local Similarity: 33.78% Mismatches: 94
Query Match: 26.07% Indels: 53
DB: 10 Gaps: 11

US-10-069-353A-8 (1-275) x US-09-953-348-36 (1-852)

QY 2 LeuProGlyGlyAlaProThrSerGlnGlnValGlyGlnMetTYRAspLeuValThrPro 21
Db 25 CTCCCATGCTCCCTCACCCTCGGCTCGAGAGATGGGCGCGCTACGACCGGTTGACGCGG 84
QY 22 LeuLeuAspSerValAlaGlyIYProCysAlaIleHisGlyTYTrpGluAsn--- 40
Db 85 CTGGAGAGCGGCTCCCTCGGAG-----AACCTGCATTCGCTGCTACTGGGACTCCGCC 138
QY 41 AspGlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
Db 139 GACAGCCAGGTGGCGCTGGCGGAGGCCACGACCGGCTCACAGCATGATGTCGAGCGG 198
QY 61 ThrValLeuAspGlyValArgLeuLeuAspValGlyCysGlyTYThrGlyInProAla 80
Db 199 CTGGCATGCGCGCGGCTCCGCGCTCTGAGCTCGGCTCGGCTGGGAGACCGCGGCGC 258
QY 81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGln 100
Db 259 GTAGCATCGGCTCGGCTCAGAGGCGGATGATCAAGGCACTCTGTGAGCATGAGAGAG 318
QY 101 ValAlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120
Db 319 GTCTCCGCGCAACGCGCTGGCGAGGAGCGCGGCTCGCGGACCGGCGCTTCGAG 378
QY 121 CysValAspAlaMetSerLeuProTYTrpProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
Db 379 CGGGCGAGCGGATGAGCTCTCCCTTGAAGACAGAGCTTGAAGCGCTCATGCGCTTC 438
QY 141 GlnSerLeuLeuGlnMetSerGluProAspArgAla-----IleArgGluIleLeuArg 158
Db 439 GAATGATCATCCACATG-----CCGACCGGCCCAAGGCTCGGCCAGAGTGGGCGG 492
QY 159 ValLeuLysProGlyGlyIleLeuGlyValThrGluValValLysArg----- 174
Db 493 GTGCTGCGGCGCGGGGCGGTCTGTGCTCACCGCATCTTTCGAGCGGGCGCCCTCGCC 552
QY 175 ---GluAlaGlyGlyGlnMetPro-----ValSer 183
Db 553 CCGAGGGGCGGGCGCGGCTCAGCGCTACCTCCACAGACTTCATGATGACCATGAGTACG 612
QY 184 GlyAspArgTrpProThrGlyLeuArg-----IleCys 194
Db 613 GCGAGGCGTACCTCCCTGCTCTCGGGGCGGCGCTGCGGAGAGATTCCTCTGAC 672
QY 195 LeuAlaGlnIleLeuLeuGlnSerLeuArgAlaAlaGlyPheGluIleLeuAspTrpGlu 214
Db 673 ATCAGCGACGAGACCTTGAGAGAGC-----TTCAGGCTGCTCTCG---GAG 717
QY 215 AspValSerSerArgThrArgTYRPhemeProGlnPheAlaGlnGluLeuAlaAlaHis 234
Db 718 CGCATCACTCTCGAAGACAGAGGCTGAGACGAGATTGCGGAGAGATG---GTGAAAC 774
```

```

Qy      235 GlnHsGlyIleAlaAspArgTyrGlyProAlaValAlaGlyTrrAlaAlaValCys 254
      |||
Db      775 CAGTTCGACCCCGCGCAGCTCTCGCGC----- 801
Qy      255 AspTyrGluLysTyrAlaHisAspMetGlyTyrAlaIleLeuThrAlaArgLysPro 273
      |||
Db      802 -----GTCAAGAGTTCCGCTATCTGCTGCTGCTGCCCGACGCCCGC 843

RESULT 9
US-10-267-255-36
Sequence 36, Application US/10267255
Publication No. US20030124689A1
GENERAL INFORMATION:
APPLICANT: Sherman, D
APPLICANT: Mao, Y
APPLICANT: Varoglu, M
APPLICANT: He, M
APPLICANT: Sheldon, P
TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
FILE REFERENCE: 600.456051
CURRENT APPLICATION NUMBER: US/10/267,255
CURRENT FILING DATE: 2002-10-09
PRIOR APPLICATION NUMBER: US 09/266,965
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: US 08/624,447
PRIOR FILING DATE: 1996-08-19
PRIOR APPLICATION NUMBER: PCT/US94/11279
PRIOR FILING DATE: 1994-10-06
PRIOR APPLICATION NUMBER: US 08/133,963
PRIOR FILING DATE: 1993-10-07
NUMBER OF SEQ ID NOS: 145
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 36
LENGTH: 852
TYPE: DNA
ORGANISM: Streptomyces lavendulae
US-10-267-255-36

Alignment Scores:
Pred. No.: 4,866-35 Length: 852
Score: 376.50 Matches: 101
Percent Similarity: 50.84% Conservative: 51
Best Local Similarity: 33.78% Mismatches: 94
Query Match: 26.07% Indels: 53
DB: 15 Gaps: 11

US-10-069-353A-8 (1-275) x US-10-267-255-36 (1-852)
Qy      2 LeuProGlyGlyAlaProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThrPro 21
      |||
Db      25 CTCGCCATGCGCCACACCGCGCTCGAGAGAGTGGCGCGCTTACGACCGGTTACCGCG 84
Qy      22 LeuLeuAsnSerValAlaGlyGlyProCysAlaIleHisGlyTyrTrpGluAsn--- 40
      |||
Db      85 CTGGAGCCCGCTCCCTCCGCGAG-----AACTGACCTTCGCTACCTGAGACTCCCCC 138
Qy      41 AspGlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
      |||
Db      139 GACAGCCAGGTGCGCGCTCGCGGACGACCGACCGGCTCACGACGATGCGCGGCGG 198
Qy      61 ThrValLeuAspGlyGlyValArgLeuLeuAspValGlyCysGlyThrGlyGlnProAla 80
      |||
Db      199 CTGCGCATCGCGCGCGCTCCCGCTCGGACTCGGCTGCGCGCTGCGGACCCCGGCG 258
Qy      81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGln 100
      |||
Db      259 GTAGGCAATCGCGCGCTCGACGAGCGCATGTCAAGGCAATCTCGGTGAGCAATGACAG 318
Qy      101 ValAlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120
      |||
Db      319 GTCTCTCGGCGCAACGCGCTGCGGAGAGCGCGGCTCGCGCAACGCGGCGGCTTCAG 378

```

```

Qy      121 CysValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
      |||
Db      379 CGGGCCGACGCGATGAGTACCTCCCTTCAGAGACGAGAGCTTGACCGCGTACCGCCCTC 438
Qy      141 GlnSerLeuLeuGlnMetSerGluProAspArgAla-----IlaArgGluIleLeuArg 158
      |||
Db      439 GAATGCATCATCATCATG-----CCGACCGCGCGCCAGGTCTCGCCCGAGGTGCGCGG 492
Qy      159 ValLeuLysProGlyGlyIleLeuGlyValThrGluValAlaValArg----- 174
      |||
Db      493 GTGCTGGCGCGCGCGCGCTGTGGTGTCTACCGACTTCTTCGACGGCGGCGCCCTCGCC 552
Qy      175 -----GluAlaGlyGlyGlyMetPro-----ValSer 183
      |||
Db      553 CCGAGAGGCGCGCGCGCGCTCCAGGCTTACCTTCACACACTTATGATGACATGATGATGAC 612
Qy      184 GlyAspArgTyrProThrGlyLeuArg-----IleCys 194
      |||
Db      613 GCGAGGCGTACCTTCCCTGCGCGGCGCGCTGTGAGCTGAGAGATTCCTCGAC 672
Qy      195 LeuAlaGluGlnLeuGlnLysLeuArgAlaAlaGlyPheGluIleLeuAspTrpGlu 214
      |||
Db      673 ATCAGGACCAAGACCTCGAGAAAGAC-----TTGAGCTGCTTCG---GAG 717
Qy      215 AspValSerSerArgThrArgTyrPheMetProGlnPheAlaGluLeuAlaAlaHis 234
      |||
Db      718 CGCATCACTCTTCGAGAGAGCGCTGAGACGCACTTCGCGGAGAGATG---GTGAC 774
Qy      235 GlnHsGlyIleAlaAspArgTyrGlyProAlaValAlaGlyTrrAlaAlaValCys 254
      |||
Db      775 CAGTTCGACCCCGCGCAGCTCTCGCGC----- 801
Qy      255 AspTyrGluLysTyrAlaHisAspMetGlyTyrAlaIleLeuThrAlaArgLysPro 273
      |||
Db      802 -----GTCAAGAGTTCCGCTATCTGCTGCTGCTGCCCGACGCCCGC 843

RESULT 10
US-09-953-348-76/C
Sequence 76, Application US/09953348
Publication No. US20030134398A1
GENERAL INFORMATION:
APPLICANT: Sherman, David. H
APPLICANT: Mao, Yingqing
APPLICANT: Varoglu, Mustafa
APPLICANT: He, Min
APPLICANT: Sheldon, Paul
TITLE OF INVENTION: MITOMYCIN BIOSYNTHETIC GENE CLUSTER
FILE REFERENCE: 600.53051
CURRENT APPLICATION NUMBER: US/09/953,348
PRIOR APPLICATION NUMBER: PCT/US00/06394
PRIOR FILING DATE: 2000-03-10
PRIOR APPLICATION NUMBER: 09/266965
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 153
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 76
LENGTH: 53500
TYPE: DNA
ORGANISM: Streptomyces lavendulae
US-09-953-348-76

Alignment Scores:
Pred. No.: 1e-32 Length: 53500
Score: 376.50 Matches: 101
Percent Similarity: 50.84% Conservative: 51
Best Local Similarity: 33.78% Mismatches: 94
Query Match: 26.07% Indels: 53
DB: 10 Gaps: 11

US-10-069-353A-8 (1-275) x US-09-953-348-76 (1-53500)
Qy      2 LeuProGlyGlyAlaProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThrPro 21

```


Qy 215 AspvAlSerSerArgThrArgTyRPhemecProGlnPheAlaGluGluLeuAlaAlaHis 234
 Db 6867 CGCATCACTCTCCGAGAGAGAGCTGAGACCGCACTTCGCGCGAGATG---GTGAAC 6811
 Qy 235 GlnHiGlyIleAlaAspArgTyRgIyProAlaValAlaGlyTrpAlaAlaAlaValCys 254
 Db 6810 CAGTTGACCCCGCGCACTTCGCGC----- 6784
 Qy 255 AspTyRgIuLeuTyRAlaHisAspMetGlyTyRAlaIleLeuThrAlaArgLysPro 273
 Db 6783 -----GTCAAGAGATTGCGCTATCTGTCTGTGCGCCGAGCCCG 6742

RESULT 12
 US-09-980-217-1/c
 ; Sequence 1, Application US/09980217
 ; Publication No. US20040219645A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Biotech Technology Limited
 ; APPLICANT: Leadlay, Peter F
 ; APPLICANT: Stanton, James
 ; APPLICANT: Olinyk, Marko
 ; TITLE OF INVENTION: Polyketides and their synthesis
 ; FILE REFERENCE: IS/bp5858469
 ; CURRENT APPLICATION NUMBER: US/09/980,217
 ; CURRENT FILING DATE: 2002-05-06
 ; PRIOR APPLICATION NUMBER: PCT/GB00/02072
 ; PRIOR FILING DATE: 2000-05-30
 ; PRIOR APPLICATION NUMBER: GB 9912563.5
 ; PRIOR FILING DATE: 1999-05-28
 ; NUMBER OF SEQ ID NOS: 52
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 30000
 ; TYPE: DNA
 ; ORGANISM: Streptomyces cinnamonensis
 ; FEATURE:
 ; OTHER INFORMATION: Nucleotides 1 to 30000 of the monensin
 ; OTHER INFORMATION: biosynthetic gene cluster
 US-09-980-217-1

Alignment Scores:
 Pred. No.: 3e-31 Length: 30000
 Score: 361.50 Matches: 90
 Percent Similarity: 50.00% Conservative: 53
 Best Local Similarity: 31.47% Mismatches: 110
 Query Match: 25.03% Indels: 33
 DB: 11 Gaps: 6

US-10-069-353a-8 (1-275) x US-09-980-217-1 (1-30000)
 Qy 6 AlaProThrSerGlnGlnValGlyGlnMetTyR-----AspLeuValThr 20
 Db 10411 GCCCCGAGCCGAGCCGAGCATTCGCGCACTACTACACACCAAGCTTCGACCTGATACT 10352
 Qy 21 ProLeuLeuSerValAlaGlyIyProCysAlaIleHisGlyTyRTrpGluAsn 40
 Db 10351 CAGCTC-----GGAGACGGCAATCTGCATCGCGCTACTGCTTGAC 10310
 Qy 41 AspGly-----ArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuAla 58
 Db 10309 GCGCGGACGAGCAGCGGCGCGCTTCGACGAGCGCATGCTCCAGACGAGATATC 10250
 Qy 59 GluArgThrValLeuAspGlyIyValArgLeuLeuAspValGlyCysGlyThrGlyGln 78
 Db 10249 CGCGGCTTCGACCGCGCGCGCGCGCGCTCTTCGACATCGGCTCGGAGAGGAC 10190
 Qy 79 ProAlaLeuArgValAlaArgAspAlaIleGlnIleThrGlyIleThrValSerGln 98
 Db 10189 CCGGCGCATGAGCTGCGCGCGCGCGCGCGCATGCGAGGTGCTCGCATCTCGTACGCGC 10130
 Qy 99 ValGlnValAlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAsp 118

Db 10129 CGCCAGCTGAGCGCGGACCCGCGCGCGCGAGCCCGCTTCGCGAGCGGCTCCG 10070
 Qy 119 PheSerCysValAspAlaMetSerLeuProTyRProAspAlaPheAspAlaAlaTrp 138
 Db 10069 TTCGAGCAGGTGAGCGCATTAAGCTTCGCTGACGACAGGCTCTTCGACCACTGCTCG 10010
 Qy 139 AlaMetGlnSerLeuLeuGluMetSerGluProAspArgAlaIleArgGluIleLeuArg 158
 Db 10009 GCCCTGAGTTCATGCTGACATGCGGACAGACAGCAGAGGTCTCACGAGGCCACCGC 9950
 Qy 159 ValLeuLysProGlyIyIleLeuGlyValThrGluValValLysArgGluAlaGlyIy 178
 Db 9949 GTCGTAAAGCCCGCGCGCGCGCGCATGCGCGCATGCTTACTCAACCCGATCC 9890
 Qy 179 GlyMetProValSerGlyAspArgTrpProThrGlyLeuArgIleCysLeuAlaGlu--- 197
 Db 9889 AGCCGCGCCCGAGCGCGCACCGCTGACGACACACAGATATACCGCGCTCACCGACATC 9830
 Qy 198 ---GlnLeuLeuGlnSerLeuArgAlaAlaGlyPheGluIleLeuAspTrpGluAspVal 216
 Db 9829 GGGGACTACCCGACATCTTCGCGCGCGCGCTGACCGCTCTGAACTGACCGACATC 9770
 Qy 217 SerSerArgThrArgTyRPhemecProGlnPheAlaGluGluLeuAlaAlaHisGlnHis 236
 Db 9769 ACCCGGAGACCGCGCAAGACTACGAGCTGATGATGATCGCGCCACCGCGGAC 9710
 Qy 237 GlyIleAlaAspArgTyRgIyProAlaValAlaGlyTrpAlaAlaAlaValCysAspTyR 256
 Db 9709 GAGTACGTGACATCATCGGC-----GTAGAGGATAC 9677
 Qy 257 GluLeuTyRAlaHis-----AspMetGlyTyRAlaIle 267
 Db 9676 GAGCTTCTCTGACCAACGAGCGCGCTTCGCGAGATGCCGAGCTCGGATCATCTTC 9617
 Qy 268 LeuThrAlaArgLysPro 273
 Db 9616 GCGACCGCGAGCGCC 9599

RESULT 13
 US-09-953-348-35
 ; Sequence 35, Application US/09953348
 ; Publication No. US20030134398A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sherman, David. H
 ; APPLICANT: Mao, Yingqing
 ; APPLICANT: Varoglu, Mustafa
 ; APPLICANT: He, Min
 ; APPLICANT: Sheldon, Paul
 ; TITLE OF INVENTION: MITOMYCIN BIOSYNTHETIC GENE CLUSTER
 ; FILE REFERENCE: 600.530US1
 ; CURRENT APPLICATION NUMBER: US/09/953,348
 ; CURRENT FILING DATE: 2001-09-12
 ; PRIOR APPLICATION NUMBER: PCT/US00/06394
 ; PRIOR FILING DATE: 2000-03-10
 ; PRIOR APPLICATION NUMBER: 09/266965
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 153
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 35
 ; LENGTH: 828
 ; TYPE: DNA
 ; ORGANISM: Streptomyces lavendulae
 US-09-953-348-35

Alignment Scores:
 Pred. No.: 2.83e-31 Length: 828
 Score: 345.00 Matches: 89
 Percent Similarity: 49.83% Conservative: 56
 Best Local Similarity: 30.58% Mismatches: 102
 Query Match: 23.89% Indels: 44
 DB: 10 Gaps: 10

US-10-069-353a-8 (1-275) x US-09-953-348-35 (1-828)

Qy	6	AlpRhrhsrserglnlgnlglgymtctYrAspbleValThrProLleuLamSer	25
Db	13	GGCGGCCCTCCGACCGGANTGTCGAACCTTCAACCGCGTCCAGCATTTATGTGTGCAC	72
Qy	26	ValAlaGlYglYProCysAlaIleHisIsglYTrTPrgLuAsn--AspGIYArgLa	44
Db	73	GGCGAAGCGGC-----TACATGCAACGGTGGCTACTGGGCGGACCGGACGTCCCAACG	126
Qy	45	SetTrpGlnIlnlAlaIAspArgLeuThrAspLeuValAlaGluArgThValLeuAsp	64
Db	127	ACGGTGAAGAAGGAGGCGACCGGCTGACCGACTACGTACTGTGGAAAGCGCTCGCGCTGCC	186
Qy	65	GIYglYValArgLeuLeuAspValGIYcysglYTrhGlnIProAlaLeuArgValAla	84
Db	187	CCGGGGGAGGGGTGGCTGCACGTGGGGTGGGGCAACGGGAAGCCACTTCGGATGCC	246
Qy	85	ArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGlnValAlaIleAla	104
Db	247	GGCCGGGACCGGGGTGGGGCCACCGGGGTCTCCATCAACCCCTACAGATGGGTCTTGC	306
Qy	105	AlaAspCysValaArgGluArgGlyLeuSerHisArgValaAspPheSerCysValaAspLa	124
Db	307	CGGCACTCCCGAGAAAGAGGCG--GACGAGGCGACCGGAGTTCCGATCGGTGAACATG	363
Qy	125	MetSerLeuProTYrProAspAsnAlaPheAspAlaAlaTrpAlaMetGlnSerLeu	144
Db	364	CTGGCGTCCCTTCTCCGACGGCTGTTGCAGCGCTGTACGCGATGGAAGATCTGC	423
Qy	145	GluMetSerGluProAspArgAlaIleArgGluIleLeuArgValLeuLysProGly	164
Db	424	CACGCCCTGGAAAGGGCGCCGACGCTTCAACCGAGATCGCCGGGTGTCTCGCCCGGGCGGC	483
Qy	165	IleLeuGlyValThrGlnValValLysArgGluAlaGlyGlyMetProValSer--	183
Db	484	CGGGTACGGTGAAGGACTTCTGTGGCGCG--CCCTGAGGAC	528
Qy	184	-----GlyAspArgTrpProThrGlyLeuArgIleCys	194
Db	529	GGCGCCGAGCATGTCGACACCGGCAACGACAACTTCAAGAGGCG--CCGCTCTC	585
Qy	195	LeuAlaGluGlnLeuLeuGlnSerLeuArgAlaAlaGlyPheGluIleLeuAspTrpGlu	214
Db	586	ACCGCGAGGCGTACGAGACTGCATGCGGTGGTGGGGCTGGAGGTGTGGAGTTCTC	645
Qy	215	AspValSerSerArgThrArgTYrPheMetProGlnPheAlaGluGlnLeuAlaHis	234
Db	646	GACATCGGGAGCAGAGTGGG-----CCCTCTAC--GAGCGGTGGCGGCAAG	693
Qy	235	GlnHisGlyIleAlaAspArgTYrGlyProAlaValAlaGlyTrpAlaAlaValCys	254
Db	694	ATGGGTGCGGCGAGGAGCGAGTGGCTCCGACATG-----	729
Qy	255	AspTYrGluLysTYr-----AlaHisAsp	262
Db	730	GAGGACGAGCGCTTCCACCGCATGTGTGACGCGGATCGACCGCATGGCTCGTGAAGAG	789
Qy	263	MetGlyTYrAlaIleLeuThrAlaArgLysPro	273
Db	790	GTCGGCTACTCGGTGTCTACCGCGCGAACC	822
RESULT 14			
US-10-267-255-35			
Sequence 35, Application US/10267255			
Publication No. US20030124689A1			
GENERAL INFORMATION:			
APPLICANT: Sherman, D			
APPLICANT: Mao, Y			
APPLICANT: Varoglu, M			
APPLICANT: He, M			
APPLICANT: Sheldon, P			
TITLE OF INVENTION: Mito-cycin biosynthetic gene cluster			
FILE REFERENCE: 600.456051			

[illegible]

```
Db      646 GACATGGGAGCAGCTGCG-----CCCTCTAC---GAGCGGTGGCGGGAAG 693
Qy      235 GlnHieGlyIleAaAspArgTyrglyProAlaValAlaGlyTrrAlaAlaValCys 254
Db      694 ATGCGGCGGCGCCAGGAGCAGCTCGGCTCCACATG-----729
Qy      255 AepTyrglyulysTyr-----AlaHisAsp 262
Db      730 GACGAGCAGCGCTTCCACCGCATGTGACGCGCATGACCGCATGGGCTCGTGAGAG 789
Qy      263 MetGlyTrrAlaIleuThrAlaArglyPro 273
Db      790 GTGCGCTACTCGGTGTGACCGCGGGAACCG 822

RESULT 15
US-10-496-377-1
/ Sequence 1, Application US/10496377
/ Publication No. US20050118590A1
/ GENERAL INFORMATION:
/ APPLICANT: Piel, Jörn
/ TITLE OF INVENTION: Novel Gene Cluster of Pederin Biosynthesis Genes
/ FILE REFERENCE: 033285-020
/ CURRENT APPLICATION NUMBER: US/10/496,377
/ PRIOR FILING DATE: 2004-05-24
/ PRIOR APPLICATION NUMBER: PCT/EP02/13085
/ PRIOR FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: EP 01127395.0
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: Patentln Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 52659
/ TYPE: DNA
/ ORGANISM: Unknown Organism
/ FEATURE:
/ OTHER INFORMATION: Description of Unknown Organism: Bacterial
/ OTHER INFORMATION: Symbiont of Paederus fuscipes
/ FEATURE:
/ OTHER INFORMATION: Sequence of the pederin biosynthetic gene cluster.
/ OTHER INFORMATION: Differences in the sequence are dependent on
/ OTHER INFORMATION: whether the DNA was sequenced from the cosmid or
/ OTHER INFORMATION: the genome
/ FEATURE:
/ NAME/KEY: gene
/ LOCATION: (274)..(1212)
/ OTHER INFORMATION: Peda putative methyltransferase
/ FEATURE:
/ NAME/KEY: gene
/ LOCATION: (1226)..(2620)
/ OTHER INFORMATION: Pedb putative FMN-dependent oxidoreductase
/ FEATURE:
/ NAME/KEY: gene
/ LOCATION: (2796)..(3800)
/ OTHER INFORMATION: Pedc putative acyltransferase
/ FEATURE:
/ NAME/KEY: gene
/ LOCATION: (3919)..(5013)
/ OTHER INFORMATION: Pedd putative acyltransferase
/ FEATURE:
/ NAME/KEY: gene
/ LOCATION: (5243)..(6064)
/ OTHER INFORMATION: Pede putative methyltransferase
/ FEATURE:
/ NAME/KEY: gene
/ LOCATION: (6309)..(32114)
/ OTHER INFORMATION: Pedf mixed type I polyketide synthase/nonribosomal
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (6305)..(8489)
/ OTHER INFORMATION: module 1 PKS
/ FEATURE:
```

```
/ NAME/KEY: misc feature
/ LOCATION: (8585)..(11618)
/ OTHER INFORMATION: module 2 NRPS
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (11817)..(15236)
/ OTHER INFORMATION: module 3 PKS
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (15381)..(20192)
/ OTHER INFORMATION: module 4 PKS
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (20334)..(25508)
/ OTHER INFORMATION: module 5 PKS
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (25686)..(29315)
/ OTHER INFORMATION: module 6 PKS
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (29478)..(32114)
/ OTHER INFORMATION: module 7 incomplete PKS
/ FEATURE:
/ NAME/KEY: gene
/ LOCATION: (32142)..(33443)
/ OTHER INFORMATION: Pedg putative flavin-binding monooxygenase
/ FEATURE:
/ NAME/KEY: gene
/ LOCATION: (33467)..(52267)
/ OTHER INFORMATION: Pedh mixed type I polyketide synthase/nonribosomal
/ OTHER INFORMATION: peptide synthetase
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (33467)..(33863)
/ OTHER INFORMATION: module 1 incomplete PKS
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (34070)..(38558)
/ OTHER INFORMATION: module 2 PKS
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (38678)..(42209)
/ OTHER INFORMATION: module 3 PKS
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (42314)..(45836)
/ OTHER INFORMATION: module 4 PKS
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (45944)..(48128)
/ OTHER INFORMATION: module 5 PKS
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (48336)..(52267)
/ OTHER INFORMATION: module 6 NRPS
/ US-10-496-377-1

Alignment Scores:
Pred. No.: 1,03e-19
Score: 268.00
Percent Similarity: 49.47%
Best Local Similarity: 28.83%
Query Match: 18.56%
DB: 21
Length: 52659
Matches: 81
Conservative: 58
Mismatch: 109
Indels: 33
Gaps: 10

US-10-069-353a-8 (1-275) x US-10-496-377-1 (1-52659)
Qy      10 GlnGlnValGlyGlnMetTyrrAp-----LeuValThrProIleuLeuAsnSer 25
Db      5267 GAGAAAGTTCCTACGCTTATATGATTCGCGAAGGCGCAAGGCGCTTATTTA-----5320
Qy      26 ValAlaGlyGlyProCysAlaIleuHisGlyTyrTrp---GluAsnAspGlyArgAla 44
```

Db 5321 ---TTCCGGCGA-----CATATGATTTGGGGCTTATTGGATGAAGTCACTGGGAGGG 53711

Qy 45 SerTrpGlnGlnAlaIaAspArgLeuThrAspLeuValAlaGluArgThrValLeuAsp 64

Db 5372 AACCTTGGCAACGCGACGAGAAAGGCTTCGCGAGATTAGATTGGCTAAGGCACTATAAAG 54311

Qy 65 G1yG1ValArgLeuLeuAspValG1yGysG1yThrG1yGlnProAlaLeuAspValAla 84

Db 5432 GCCGGGCAAAATTTATCATGAGATGGAGATGGCTTTGGAGAGGTCTGGCTTAAAGCTGGC 54911

Qy 85 ArgAspAsnAlaIleGlnIleThrG1yIleThrValSerGlnValGlnValAlaIleAla 104

Db 5492 AAAGCCAAAGGGTGTGTTGTTGATGAGATTAAACATTATACAAAGAGCAACACTCAGTGA 55511

Qy 105 AlaAspCySAlaArgGluArgG1yLeuSerHisArgValAspPheSerCysValAspAla 124

Db 5552 ATTACTCGGGGCTGAAGCAGACAGCAACTGCAAGAGCGTGCATCTATTATTCATGTAGCGCA 56111

Qy 125 MetSerLeuProTyIyProAspAsnAlaPheAspAlaAlaITrpAlaMetGlnSerLeuLeu 144

Db 5612 CTGAATATACCGTGGAGAACCAATCGATATGATGTGTGGTTTTCGATCACTATTTT 56711

Qy 145 GluMetSerGluProAspArgAlaIleArgGlnIleLeuArgValLeuLysProG1yG1y 164

Db 5672 CATATGGGTCA--AGAAAGCCCTTCATGAGGCTGCGCGAGTCACTCAACCTGGTCT 57281

Qy 165 IleLeuG1yValThrGluValValLysArgGluValG1yG1yGlyMetProValSerG1y 184

Db 5729 ACTTATTATTAAACGATTGTGCTTT-----CTTCTCGAGAGTACC 57701

Qy 185 AspArg-----TrpPro-ThrG1yLeuArgIle-CysLeuAlaGluGln 199

Db 5771 GAAGCTTTTAAGAGAGTCTTTGGGAACACATTCATTGCGCTTTTGTCTCGGAGAGAT 58301

Qy 199 eulEu-glusErLeuArgAlaIaIaG1yPheGluIleLeuAspTrpGluAspValSerSer 218

Db 5831 TATCTTCGTGCTATTAAAGCCGAAGCGGAGTTCGAACATCGAAGAAATGATGATTATTCGAT 58901

Qy 219 ArgThrArg--TyrPheMetProGlnPheAlaGluGlnLysuAlaAlaHisGlnHisGly 237

Db 5891 AATGTCATGCCCTGGCTGGAGGCCCAAGTTGAAAGAGAGATCGAATCGATCGAACCGCAG 59501

Qy 238 -----IleAlaAspArgTyG1yProAlaValAlaG1yTrpAlaAlaAlaVal 253

Db 5951 GTCCGAGCGCATATTCTCTAATGATACGAGAAGGAGATCGACGATGGCTCTATCTG-- 60071

Qy 254 CysAspArgTyG1yLysTyAlaHisAspMetG1yTyAlaIleLeuThrAlaArgLys 272

Db 6008 -----TTCCAGTACATGAGGAGAACTTAAAGCTTACATGATGTGGATGGCTAAAG 60581

Search completed: August 29, 2005, 23:46:40
Job time : 5041 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 29, 2005, 21:22:55 ; Search time 4680 Seconds
(without alignments)
2847.262 Million cell updates/sec

Title: US-10-069-353A-8

Perfect score: 275
Sequence: 1 VLBGAPTSQVQNGMYDVT.....YKXAHDMGVALITAPKPVG 275

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 1

Total number of hits satisfying chosen parameters: 9408497

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O/cg12.1/USPTO.spool/US10069353/runat.26082005.172151.15052/app.query.fasta_1.455
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MIMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10069353 @CG1_1.1 3731 @runat.26082005.172151.15052 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBIOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	275	100.0	828	6	AX089422 Sequence
2	275	100.0	45624	6	AX089419 Sequence
3	275	100.0	50000	6	AX089416 Sequence
4	275	100.0	80161	1	AY007564 Saccharop

5	275	100.0	80161	6	AR165018 Sequence
6	275	100.0	80161	6	BD137649 Biosynthe
7	275	100.0	80161	6	AR281866 Sequence
8	37	13.5	36538	6	AX600587 Sequence
9	10	3.6	738	6	AR534398 Sequence
10	10	3.6	738	6	AX573660 Sequence
11	10	3.6	788	8	BT009270 Triflicum
12	10	3.6	1394	5	AY491379 Acanthop
13	10	3.6	6085	6	AE0304
14	10	3.6	6085	6	AR144763 Sequence
15	10	3.6	10802	1	AE005764
16	10	3.6	22399	1	AB088119 Streptom
17	10	3.6	90348	1	AF497482 Microomo
18	10	3.6	297850	1	AP006577 Gloeobact
19	10	3.6	298550	1	AP005961 Bradyrhiz
20	10	3.6	298750	1	AP005375 Thermosyn
21	9	3.3	400	6	AR356861 Sequence
22	9	3.3	400	6	AR358417 Sequence
23	9	3.3	759	6	AK619480 Sequence
24	9	3.3	2191	5	AB034983 Xenopus 1
25	9	3.3	10993	1	AE011153 Methanosa
26	9	3.3	11466	6	AR354326 Sequence
27	9	3.3	11466	6	AR355882 Sequence
28	9	3.3	50049	1	AP003358 3
29	9	3.3	54587	2	AC146908_3
30	9	3.3	64492	1	AB086653 Streptomy
31	9	3.3	66201	1	AB003601 Nostoc sp
32	9	3.3	95684	2	AC134115 Rattus no
33	9	3.3	98835	9	AC073195 Homo sapi
34	9	3.3	110000	1	AB017180_16
35	9	3.3	110000	1	BX571856_02
36	9	3.3	110000	1	BX571856_03
37	9	3.3	110000	1	BX571857_02
38	9	3.3	13281	8	AC147802 Oryza bat
39	9	3.3	152950	2	AC121067 Rattus no
40	9	3.3	160031	4	AC147681 Canis Fam
41	9	3.3	164398	10	AC116808 Mus muscu
42	9	3.3	169065	2	AP002338 Homo sapi
43	9	3.3	201472	2	AC073924 Homo sapi
44	9	3.3	216172	9	AC098973 Homo sapi
45	9	3.3	223337	10	AL928564 Mouse DNA

ALIGNMENTS

RESULT 1
AX089422 828 bp DNA linear PAT 21-MAR-2001
LOCUS Sequence 7 from Patent WO0116303.
DEFINITION AX089422
ACCESSION AX089422
VERSION GI:13443683
KEYWORDS
SOURCE
ORGANISM
Saccharopolyspora spinosa
Saccharopolyspora spinosa
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Pseudonocardineae; Pseudonocardiaceae; Saccharopolyspora.
REFERENCE
Eberz, G., Moehle, V., Froede, R., Velten, R. and Salas, V.A.
Nucleic acids which code for the enzyme activities of the spinosyn
biosynthesis
JOURNAL
Patent: WO 0116303-A 7 08-MAR-2001;
BAYER AG (DE)
FEATURES
source
1..828
Location/Qualifiers
/organism="Saccharopolyspora spinosa"
/mol_type="unassigned DNA"
/db_xref="taxon:60894"
1..828
/note="unnamed protein product; ORF1, O-Methyltransferase"

/codon_start=1
/transl_table=11
/protein_id="CAC34874.1"
/db_xref="GI:13443684"

/translation="MLPGAPTSQGVQWYDVTLPPLNSVAGSPCAIHGHWENDGRA
SMOQADRLTDLVARTVLDGVRILDGCGTGPALPAPARADAIOTIGTIVSOYVA
IADACRERGLSHRDVPSVDAMSILPYPNAPRPAAMSLMSPPDAIETILPVL
KRGILGTVBTVKREAGGMPVSGDBWPGRLCLAEQLLESIRAAQFEIILMEDVSS
RRTYFMPQAEELAHQHGLADRYGPVAVGMAAAVCDYKRYADHWGYALTTPAKPVG"

ORIGIN

Alignment Scores:

Pred. No.:	4,63e-292	Length:	828
Score:	275.00	Matches:	275
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-069-353a-8 (1-275) x AX089422 (1-828)

QY	1	ValLeuProGlyValAlaProThSerGlnGlnValGlyGlnMetTyAspLeuValThr	20
DB	1	GTGTTCCAGTGGCGCACCAACATCGCAGAGTTGGGAGATGATGACTGTGCAG	60
QY	21	ProLeuLeuAsnSerValAlaGlyProCysAlaIleHisGlyTyTrpGluAsn	40
DB	61	CGGTGCTGAACTCGGTCGGGGGCGCCCTCGCCATCCACAGGCTACGCGAGAAC	120
QY	41	AspGlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg	60
DB	121	GACGGCGGGCTTCCTGGCAGCAGCGCCGCAACGGCTCACCTTGTCCCGCAACG	180
QY	61	ThrValLeuAspGlyValAlaArgLeuLeuAspValGlyCysGlyThrGlyGlnProAla	80
DB	181	ACCGTCTCGATGGCGCGCTTCGACTGCTGATGTGGGGTGCGGTACCGGACCAACAGCG	240
QY	81	LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyTyleThrValSerGlnValGln	100
DB	241	CTGGCGGTGGCGGCAACAGCGATCCAGATCACCGGCACTCCGTACAGCCAGGTGCA	300
QY	101	ValAlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSer	120
DB	301	GTGGCCATCGCGCTGATTCGCGCAGCAACCGGCACTAAGCAGCGGGTGACTTCTCG	360
QY	121	CysValAspAlaMetSerLeuProTyTrpAspAsnAlaPheAspAlaAlaTrpAlaMet	140
DB	361	TGGTGATGATGCAATGTCCTCGCGTACCGGCAATGCTTTCGACCGCGCTGGGCGCATG	420
QY	141	GlnSerLeuLeuGlnMetSerGluProAspArgAlaIleArgGlnIleLeuArgValLeu	160
DB	421	CAGTCCCTGTTGAGATGTCGCCAACCGGACCTGCCATCCGGAAATCTTGCAGTACTC	480
QY	161	LysProGlyGlyIleLeuGlyValThrGluValAlaArgGluAlaGlyGlyMet	180
DB	481	AAACCCGGTGGCATCTCCGGCGTCACGAGGTGTCAACGAGAACGGGGCGCGGAGATG	540
QY	181	ProValSerGlyAspArgTrpProThrGlyLeuArgIleCysLeuAlaGlnLeuLeu	200
DB	541	CCGGTGTCCGGGACGATGCGCGACCGGCTTCGATCTGCTGCTAGCAACTTCTCG	600
QY	201	GlnSerLeuArgAlaAlaGlyPheGlnIleLeuAspTrpGluAspValSerSerArgThr	220
DB	601	GAATCCCTGCTGACAGGGGGGTTCCGATCTCGATTTGGAGAGCACTGTGTCGAGAAC	660
QY	221	ArgTyTrPheMetProGlnPheAlaGlnIleLeuAlaAlaHisGlnHisGlyIleAlaAsp	240
DB	661	CGGTACTTATGCGCAGTTCGCGCAAGAGCTCGCTGCCGACACGAGCGGGATCGCGAC	720
QY	241	ArgTyTrGlyProAlaValAlaGlyTyTrpAlaAlaAlaValCysAspTyTrGluTyTrAla	260
DB	721	AGGTAGGGCGCGCTGTGCGCGCTGGCGCGCGCTGCTGCAATTATAGAAATATATGCC	780
QY	261	HisAspMetGlyTyTrAlaIleLeuThrAlaArgLysProValGly	275
DB	781	CACGACATGGGCTATGCGATTCTGACGCGCGGAGACCGGCTCGGC	825

RESULT 2
AX089419/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
Patent: WO 0116303-A 4 08-MAR-2001;

FEATURES

source	location/Qualifiers
	1..45624
	/organism="Saccharopolyspora spinosa"
	/mol_type="unassigned DNA"
	/db_xref="taxon:60894"

ORIGIN

Alignment Scores:

Pred. No.:	1.73e-290	Length:	45624
Score:	275.00	Matches:	275
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-069-353a-8 (1-275) x AX089419 (1-45624)

QY	1	ValLeuProGlyValAlaProThSerGlnGlnValGlyGlnMetTyAspLeuValThr	20
DB	828	GTGTTCCAGTGGCGCACCAACATCGCAGAGTTGGGAGATGATGACTGTGCAG	769
QY	21	ProLeuLeuAsnSerValAlaGlyProCysAlaIleHisGlyTyTrpGluAsn	40
DB	768	CGGTGCTGAACTCGGTCGGGGGCGCCCTCGGCCATCCACAGGCTACTGGAGAAC	709
QY	41	AspGlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg	60
DB	708	GACGGCGGGCTTCCTGGCAGCAGCGCCGCAACGGCTCACCTTGTCCCGCAACG	649
QY	61	ThrValLeuAspGlyValAlaArgLeuLeuAspValGlyCysGlyThrGlyGlnProAla	80
DB	648	ACCGTCTCGATGGCGCGCTTCGACTGCTGATGTGGGGTGCGGTACCGGACCAACAGCG	589
QY	81	LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyTyleThrValSerGlnValGln	100
DB	588	CTGGCGGTGGCGGCAACAGCGATCCAGATCACCGGCACTCCGTACAGCGAGGTGCA	529
QY	101	ValAlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSer	120
DB	528	GTGGCCATCGCGCTGATTCGCGCAGCAACCGGACTTAAGCAGCGGGTGACTTCTCG	469
QY	121	CysValAspAlaMetSerLeuProTyTrpAspAsnAlaPheAspAlaAlaTrpAlaMet	140
DB	468	TGGTGATGATGCAATGTCCTCGCGTACCGGCAATGCTTTCGACCGCGCTGGGCGCATG	409
QY	141	GlnSerLeuLeuGlnMetSerGluProAspArgAlaIleArgGlnIleLeuArgValLeu	160
DB	408	CAGTCCCTGTTGAGATGTCGCCAACCGGACCTGCCATCCGGAAATCTTGCAGTACTC	349
QY	161	LysProGlyGlyIleLeuGlyValThrGluValAlaArgGluAlaGlyGlyMet	180
DB	348	AAACCCGGTGGCATCTCGCGGTCAACGAGGTGTCAACGAGAACGGGGCGGGGATG	289
QY	181	ProValSerGlyAspArgTrpProThrGlyLeuArgIleCysLeuAlaGlnLeuLeu	200
DB	288	CCGGTGTCCGGGACGATGCGCGCTTCGATCTGCTGCTAGCAACTTCTGAGCAACTTCTG	229

QY 201 GluSerLeuArgAlaAlaGlyPheGluIleLeuAspTrpGluAspValSerSerArgThr 220
 Db 228 GAATCGCTGCTGACGGGGGTTTCAGATCTCGATTGGAGAGACGGTGTCTGAGAGACC 169
 QY 221 ArgTyrPheMetProGlnPheAlaGluGluLeuAlaAlaHisGlnHisGlyIleAlaAsp 240
 Db 168 CGGTACTTCAATGCCAGTTCGGCGAAGAGCTGCTGCGCAACGACACGGGATCGCGGAC 109
 QY 241 ArgTyrGlyProAlaValAlaGlyTyrPAlaAlaAlaValCysAspTyrGluValYTrpAla 260
 Db 108 AGGTACGGGCGCGCTGTGCGCGCTGGCGCGCGCGCTGCGCATTTATGAGAAATATGCC 49
 QY 261 HisAspMetGlyTyrAlaAlaIleLeuThrAlaArgLysProValGly 275
 Db 48 CACGACATGGGCTATGCGATTCTGACGCGCGGAGCGGCTCGGC 4
 RESULT 3
 AX089416 AX089416 50000 bp DNA linear PAT 21-MAR-2001
 DEFINITION Sequence 1 from Patent WO0116303.
 ACCESSION AX089416
 VERSION AX089416.1 GI:13443677
 KEYWORDS
 SOURCE
 ORGANISM
 Saccharopolyspora spinosa
 Saccharopolyspora spinosa
 Bacteria; Actinobacteriia; Actinobacteridae; Actinomycetales;
 Pseudonocardineae; Pseudonocardiaceae; Saccharopolyspora.
 REFERENCE
 AUTHORS Beyer, G., Moehrle, V., Froede, R., Velten, R. and Salaas, J. A.
 TITLE Nucleic acids which code for the enzyme activities of the spinosyn
 biosynthesis
 JOURNAL Patent: WO 0116303-A 1 08-MAR-2001;
 BAYER AG (DE)
 FEATURES
 source 1..50000
 location/Qualifiers
 /organism="Saccharopolyspora spinosa"
 /mol_type="unassigned DNA"
 /db_xref="taxon:60894"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1 88e-290 Length: 50000
 Score: 275.00 Matches: 275
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-10-069-353a-8 (1-275) x AX089416 (1-50000)
 QY 1 ValLeuProGlyAlaAlaProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThr 20
 Db 44797 GTGTGGCAGGTGGCGGACCAACATCGCAGAGTTGGCAGATGTATGACTCTGTCACG 44856
 QY 21 ProLeuLeuAsnSerValAlaAlaGlyGlyProCysAlaIleHisHisGlyTyrTrpGluAsn 40
 Db 44857 CCGTTGTGAACCTGCGCGGGCGGCGCCCTGTGCGCATTCACACACGCGCTACTGGAGAAC 44916
 QY 41 AsnGlyValArgLysSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
 Db 44917 GACGGGCGGCGCTTCTTGCGACGAGCGCGGACCGGCTCCACCACTTGTGCGCGAAGCG 44976
 QY 61 ThrValLeuAspGlyGlyValAlaTrpLeuLeuAspValGlyCysGlyTyrTrpGluProAla 80
 Db 44977 ACCGTGCTCATGGCGGCGGCTTCACATCTCGATGTGGGGTGGCGATACCGGACCAACAGCG 45036
 QY 81 LeuArgValAlaArgAspAsnAlaAlaGlnIleThrGlyValLeuThrValSerGlnValGln 100
 Db 45037 CTGGCGCTTCGCGCGGCAACCGCATCCAGATCCCGGACATCCGTCACCGCAGCGAGTCA 45096
 QY 101 ValAlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120

Db 45097 GTGGCCATGCGCGCTGATTGGCGACGCGGACGACTTAAGCCACCGGCTGACTTCTCG 45156
 QY 121 CysValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
 Db 45157 TCGGTGAGATGCGATGCTTCCGTGCGTACCGGACCAATGCTTTGACGCCCTCGGGCCATG 45216
 QY 141 GlnSerLeuLeuGlnMetSerGluProAspArgAlaAlaArgGlnIleLeuArgValLeu 160
 Db 45217 CAGTCGCTTTGAGATGTGCCAAGCGGACCGGTGCATCCGGGAAATCCTTGAGATATTC 45276
 QY 161 LysProGlyGlyIleLeuGlyValThrGluValAlaLysArgGluAlaGlyGlyMet 180
 Db 45277 AAACCGGTGGCATCTCTGGCGCTCAACGAGGTGCTCAAAAGAGAGCGGCGCGCGGAGT 45336
 QY 181 ProValSerGlyAspArgTrpProThrGlyLeuAlaGlyLeuLeuAlaGluIleLeuLeu 200
 Db 45337 CCGGTGTCGGGGACAGGTGGCGCGACCGGCTTCGATCTGCTGCTGAGCAACTTCTG 45396
 QY 201 GluSerLeuArgAlaAlaGlyPheGluIleLeuAspTrpGluAspValSerSerArgThr 220
 Db 45397 GAATCGCTGCTGACGGGGGTTTCAGATCTCGATTGGAGAGACGTGTCTGAGAGACC 45456
 QY 221 ArgTyrPheMetProGlnPheAlaGluGluLeuAlaAlaHisGlnHisGlyIleAlaAsp 240
 Db 45457 CGGTACTTCAATGCCAGATTCGCCCGAAGAGCTCGCTGGCGCACACGACGCGGATCGCGAC 45516
 QY 241 ArgTyrGlyProAlaValAlaGlyTyrPAlaAlaAlaValCysAspTyrGluValYTrpAla 260
 Db 45517 AGGTACGGGCGCGCTGTGCGCGCTGGCGCGCGCGCTGCGCATTTATGAGAAATATGCC 45576
 QY 261 HisAspMetGlyTyrAlaAlaIleLeuThrAlaArgLysProValGly 275
 Db 45577 CACGACATGGGCTATGCGATTCTGACGCGCGGAGCGGATCGGC 45621
 RESULT 4
 AY007564 AY007564 80161 bp DNA linear BCT 16-MAY-2001
 DEFINITION Saccharopolyspora spinosa probable 1y8r-type transcriptional
 regulator, probable keto acyl reductase, spinosed biosynthetic gene
 cluster, complete sequence, and probable exodeoxyribonuclease V
 genes, complete cds; and unknown gene.
 ACCESSION AY007564.1 GI:13162633
 VERSION AY007564
 KEYWORDS
 SOURCE
 ORGANISM
 Saccharopolyspora spinosa
 Saccharopolyspora spinosa
 Bacteria; Actinobacteriia; Actinobacteridae; Actinomycetales;
 Pseudonocardineae; Pseudonocardiaceae; Saccharopolyspora.
 REFERENCE
 AUTHORS Waldron, C., Matsushina, P., Rostek, P. R. Jr., Broughton, M. C.,
 Turner, J., Madhuri, K., Crawford, K. P., Merlo, D. J. and Baltes, R. H.
 TITLE Cloning and analysis of the spinosed biosynthetic gene cluster of
 Saccharopolyspora spinosa
 JOURNAL Chem. Biol. 8 (5), 487-499 (2001)
 MEDLINE 21257765
 PUBMED 11358695
 REFERENCE
 2 (bases 1 to 80161)
 Turner, J., Madhuri, K., Crawford, K. P., Merlo, D. J. and Baltes, R. H.
 TITLE Direct Substitution
 JOURNAL Submitted (18-AUG-2000) Dow Agrosciences, 9330 Zionsville Rd.,
 Indianapolis, IN 46268, USA
 FEATURES
 source 1..80161
 location/Qualifiers
 /organism="Saccharopolyspora spinosa"
 /mol_type="genomic DNA"
 /db_xref="taxon:60894"
 complement(1135..1971)
 /note="ORF L16"
 /codon_start=1
 /transl_table=11
 /product="probable 1y8r-type transcriptional regulator"
 /protein_id="AAG23282.1"

```

/db_xref="GI:13162654"
/translation="METREILRYFAVAEEELHPCRAOQLIAOPLSRTIAOLEOELG
VYLORTSRKVSITEAGMLITTEGRALILGALIAEAORLATSOSPILYAKAGSG
ELLAKLIDVAAPGAVAVDILLCESQPOKTHIDGADYALILHQPDPFAELDIETLN
TEQOVAILPTSHPLASEPHVMADVSLDPLPLARWPGDYTPDGVEVKNQOLF
QMIALGRTIVMESSRVNLLBGLAAPPVLDAPDYTVIAMPHSRSRALAGLVVAT
LI"
2024. .2791
/feature="ORF 115"
/codon_start=1
/translation="probable keto acyl reductase"
/protein_id="AAG23281.1"
/db_xref="GI:13162653"
/translation="WLVPRPRPRLATLESMEQOTIALVTGANKGIGYEIAAGCALGMS
VIGARDHORGEDAVAKLRADGVDAFVSLDVTDDASVAAAALLLEERAGRLDIVLNN
AGIAGWPREPSTVTPASIRAYETTVISVYVTAAMPLILKRSERPRVYNSSHYAS
LTHQTPGVLDGIGISGAYSPSKTFILAITTQYAKELSDTNIKINNACPGYVATDLNGF
HGTSTPADARIALRLATLPPDGPFGMPDAGNVPM"
complement(3416. .4165)
/gene="spnS"
complement(3416. .4165)
/feature="spnS"
/feature="probably involved in forosamine biosynthesis"
/codon_start=1
/evidence=not experimental
/translation_table=11
/product="probable N-dimethyltransferase"
/protein_id="AAG23280.1"
/db_xref="GI:13162652"
/translation="MSRVSDTPEITSSVYSPPHADIYDAIHSARGDMAAAGEVVOVL
VTRLRPEASGLDVAACGTGAHLERFPAEYAKVAGLELSDAMEIARIVPEVPIHGD
IHDFDLGEPPDVITCLCTFAAYKRVITDDLRVTRMNAHLAAGVAVIEPMPPKFI
DEGVTVAAHNGEGRVSRLSHVSILEGRTSRMTVRYVAEPTGIRDPTEIELSLFTEED
EXTALIEDAGIAEVLPGAPNGRGLFVGIRN"
complement(4168. .5325)
/gene="spnR"
complement(4168. .5325)
/feature="spnR"
/feature="probably involved in forosamine biosynthesis"
/codon_start=1
/evidence=not experimental
/translation_table=11
/product="probable aminotransferase"
/protein_id="AAG23279.1"
/db_xref="GI:13162651"
/translation="MINIHOPIITGTEELDAIAEVPASNWIIGPPTRTTEAPFAHILG
VDBEYVPLANSGRNAELTVQVLDLGRGDVTLPSISFYAANAATASSGARVFCVDV
PRTLNTLDVAPATITPATKAVULHLYGSGPESVTHIADFCRKGIMLTEDSACAVAS
SAGSTACGTFFGDLATWSFDAMKILVTGDCGMFYAADPELAHRARILAYGLQMSGFD
SAKSSNRWDIRVEDIGORLIGNDWTALGVSQVLRKLPFIRNRREIAYOYRLSDV
PGVLLPPTLPDGVSSHVYVQVLAPEIRDOVAQOMLERGITSYRYPPLHLYPYRA
DCKLPBAEDACRITLLPLPHSPDLDAEVTVADDFGKAVHHIISGRSPLRK"
complement(5363. .6751)
/gene="spnQ"
complement(5363. .6751)
/feature="spnQ"
/feature="probably involved in forosamine biosynthesis"
/codon_start=1
/evidence=not experimental
/translation_table=11
/product="probable NDP-hexose-3,4-dehydratase"
/protein_id="AAG23278.1"
/db_xref="GI:13162650"
/translation="MOSKRTALGKGRARVTCDDTCATATEMVPDAKDRILASVRYD
HREBSPTFVAGSTPIRSGAVLDEDDRYALVEALIELIAGGNAPRRESFAFPG
LKAHLVNSGSSANLLASLITSPKIGERLAPGDEVITTAAGPPTTINPAHQLFVEND
VAVDELGTYNATPDRIKAAVETKRALMLATLHNPFAADSIABIAKHEHLIPEJEDN
CDAVSGTYGRLLGTFGDLTVSFPYAHNITSGEGGCYVLTGIELARIIIESLDWGRD
CWCEPVDNTRCKRFDYHLGTLPGYDHKVTSGVGNKXTDLOALALISQDKISA
FSASRRMRRLREGSLGPGLLPVATPHSDPSMWGFAITTSADAGFTRALVNFLE
SPNIGRLLFGNITRHPAFEGVYRRIADALNSDLYDTRTFWGVYREITDQMDIVY
VSIABFVAKSS"

```

```

7083. .8450
/gene="spnR"
7083. .8450
/feature="spnR"
/feature="involved in forosamine addition"
/codon_start=1
/translation_table=11
/product="probable NDP-forosamy1transferase"
/protein_id="AAG23277.1"
/db_xref="GI:13162649"
/translation="WILGMLPGCSIAIEENRVLFTPLPASHFFPNVLPLMALPAAG
HEVRVAILCPNVASVVTGAGLTAVPVGDLDLISLAAKVELVSGSVSEDEGRHELF
DELSINGRDTDAVEQLHLVDVDSLDLMEFAEKMOQDLYVMDMAGSGSVVVRALG
ARHVMYALVADPSGWLIRSGFLEQSGSKPEQORVDPGLWIAKAKAFGATDEELVTG
QATIDPISMMRLPVLDLIDYISMRFPVYGPVPLPEMLREERTKRCVCTRGLTRRLS
RVTREYGSOSIOEOAMVRRLRGARLDVEYIATLSDENREMELEPNNRVYVPL
NLELSCSVITIHGSTTTOETATYANGVQLLPGLTFNDESRABELADRGGLVADPA
TFTEDVVGQARLIDEPSFANALALIREIIESSPSPHDIVPRLEKVAERENRTGQ
SDGHP"
8967. .10427
/gene="spnO"
8967. .10427
/feature="spnO"
/feature="involved in forosamine biosynthesis"
/codon_start=1
/translation_table=11
/product="probable NDP-hexose-2,3-dehydratase"
/protein_id="AAG23276.1"
/db_xref="GI:13162648"
/translation="WSSSVBEBASNAAPLGSNNTRFPVDSLALSCNMIPTTEHCWL
ADRIGENSFETNRIPEPRLSKWKCDASTENIYVHAGREFTEGLOVETNYGAAPSMQ
PILNOAGEIIGIIVKEIDGVILCHMSAKMPPGNVNVLTQSPVQATRSNTYQAHRS
VPPYVADYVNGRGLVADVLVOSQSGSWFYKRRNMVVEQEEVPLVPCMLTGLQ
VIALRQONTVMOTRVTLSGIPRDSATGPBELAASEEPQAAVARSLSHGIDSSSIS
EAVGPEFEAKARVLRATRVPILSVDDKVTPTETIADODGVPVAVIVASVATREVA
SNTQPMITPEROGETIALIVKRIIGVYLHGELHARVBAQTKYAEIAPTYQCSVANYOST
PSNDMPPLDVLVLTADPEYRESILSEEGRFYQAQNRVRIIEVHEDFAARPSDFR
MWTLGQLELRSHTFIMIQARSIVASIHSLMALGR"
10436. .11434
/gene="spnM"
10436. .11434
/feature="spnM"
/feature="involved in forosamine biosynthesis"
/codon_start=1
/translation_table=11
/product="probable NDP-hexose-3-ketoreductase"
/protein_id="AAG23275.1"
/db_xref="GI:13162647"
/translation="WRKPVRLGVLCGSAFAMRMLPAMCDVAETEVAVASRDPAAKE
RFAAFEBCEAVLGYORLLERPDIDAVVPLPPGNMAEWIIGALBAKHVLAKEPLTTP
ASDARLVEGLARKNLILRENYLPLAHGRHVVNDLQSGEIGELREFTAVFGIPLP
DTRIRYRELGGALDIDGVYPAARARIFLGLPTVAGSIAESHOESGVDGSLVLDQ
SEGGVAVLIGYFVHNHVSAYELNGSRRIRIVDAFPBPAAWQVAVIERGGVVDLS
LPABDQVKATYAFARDIRACTGYDDPRAVDAGDSGSMIQQALYBALIQARCGST"
complement(11530. .12492)
/gene="spnM"
complement(11530. .12492)
/feature="spnM"
/feature="involved in epiinosyn aglycone biosynthesis"
/codon_start=1
/translation_table=11
/product="spnM"
/protein_id="AAG23274.1"
/db_xref="GI:13162646"
/translation="MPNANSGTVLVNPIPMFREDPIITFAVTHGIGSOVAPSYLIR
TGTBETELIALVADRGVAIVITTYEGIGTCTTYTVYGRQAGMADAAAGORLPG
SGLTITDCVGIWGIQGGQASAFGEHLHTAPBLRKAALAGAVPIDULLTHENDG
VFTGVGLVGLGHAAYEPDLFPDELTLEAGTAVDQVEIARAPLVTFRFLRELSDFL
DTSGLFEQPRRARILAESVAGRNAGPVPTLVYSTDEIYPPFGEGRIDSRYAAGT
FVRNHPISGLHAPFAALAESVWVSWPFDEHFSBSAISGPADAR"
complement(12696. .13547)
/gene="spnL"
complement(12696. .13547)

```

```
/gene="spnL"
/notes="involved in spinosyn aglycone biosynthesis;
possibly involved in polyketide bridging"
/codon_start=1
/translation="methyltransferase-like protein"
/product="methyltransferase-like protein"
/protein_id="AAG3273.1"
/db_xref="GI:13162645"
/translation="MESITPDALAHGRPLHHGTYAGCYREDACATPMSDAADQTDLFI
DKALRPGALFDLGGGNGQPVRAACASVAVTGATVNAQHLAAATRLANETGLAS
LEFDLVDGQLPYDGFQAAAMQSVQVVDQAAIREVHRIIBGGRVLDIITR
VRLPEEYAAVMTGTATLINSFTALVSAFEIILEVDTAOTRCWMSVYDELRLK
DELAGEPAAVGTYQORVLDIAAKGPGPAQILAAVAEYRKHPDYARNESMGFMLL
QARKOS"
complement(13592..14785)
/gene="spnK"
complement(13592..14785)

gene
CDS
complement(13592..14785)

Alignment Scores:
Pred. No.: 2,88e-290 Length: 80161
Score: 275.00 Matches: 275
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-10-069-353a-8 (1-275) x AY007564 (1-80161)

Qy 1 ValLeuProGlyValAlaProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThr 20
Db 20168 GTGTTGCCAGTGGCGGACCAACATCGCAGAGTTGGGAGATGTATGACTGGTCAAG 20227

Qy 21 ProLeuLeuAsnSerValAlaGlyGlyProCybAlaIleHisGlyTyrTrpGluAsn 40
Db 20228 CCGTTGCTGAACCTCGGTCGGGGGGGCGCCCTGCGCCATCCACACCGGCTACTGGAGAAC 20287

Qy 41 AspGlyAArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
Db 20288 GACGGGGGGGCTTCCTGCGAGCAGCGCCGACCGGCTGACCGACTTGTGCGCGAAGCG 20347

Qy 61 ThrValLeuAspGlyValAlaArgLeuLeuAspValGlyCybGlyThrGlyGlnProAla 80
Db 20348 ACCGTGCTGATGGCGGCGGCTTCGACTGCTCGATGTGGGGTGGCGTACCGGACACCAAGCG 20407

Qy 81 LeuArgValAlaAArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGln 100
Db 20408 CTCGGCGTCCGGCGGCAACCGGATCCAGATCACCGGCTACCGCTCACCGGCTGAGCA 20467

Qy 101 ValAlaIleAlaAlaAspCybAlaAArgGluArgGlyLeuSerHisArgValAspPheSer 120
Db 20468 GTGGCCATCGCCGCTGATGCGGACCGGAGCGGAGCTAAGCCACCGGAGACTTCTCG 20527

Qy 121 CysValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
Db 20528 TGGGTGATGATCCATGTCCTCGCGGTAACCGGACATCTTTCAGCGCGCTGGGCATG 20587

Qy 141 GlnSerLeuLeuGlnMetSerGluProAspArgAlaIleArgGluIleLeuArgValLeu 160
Db 20588 CAGTCCGCTGTGAGATGTCGGAACCGGACCGGAGCTAAGCAATCTTTCAGATCACTC 20647

Qy 161 LysProGlyGlyIleLeuGlyValThrGluValAlaValAspArgGluAlaGlyGlyMet 180
Db 20648 AAACCCGGTGGCATCTCGCGGTCACCGAGGTGCTCAACGAGAGGGGGGGGGAGT 20707

Qy 181 ProValSerGlyAspArgTrpProThrGlyLeuArgIleCybLeuAlaGlnGlnLeuLeu 200
Db 20708 CCGGTGTCGGGGGACAGGTGGCGGACCGGCTTCGATCGCTGGTGGAGCACTTCTG 20767

Qy 201 GluSerLeuArgAlaAlaGlyPheGluIleLeuAspTrpGluAspValSerSerArgThr 220
Db 20768 GAATCGCTGCTGAGCGGGGTTTCGAGATCTTCGATGGAGGACGCTGTGTGAGAGACC 20827

Qy 221 ArgTyrPheMetProGlnPheAlaGluGluLeuAlaIleHisGlnHisGlyIleAlaAsp 240
```

```
Db 20828 CGGTACTTCATGCCGAGCTTCGCCGAGAGGCTGCTGCGGACCAAGACCGGATCCGGGAC 20887

Qy 241 ArgTyrGlyProAlaValAlaGlyTyrAlaAlaAlaValCybAspTyrGluYsTyrAla 260
Db 20888 AGGTACGGGCGGCTGCTCCCGGCTGGGGCGGCGCGCTGCTGCGATATAGAAATATGCC 20947

Qy 261 HisAspMetGlyTyrAlaIleLeuThrAlaArgValProValGly 275
Db 20948 CACGACATGAGGCTATGATGATTCGACGCGCGGAGACCGGATCGGC 20992

RESULT 5
LOCUS AR165018 80161 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6274350.
ACCESSION AR165018
VERSION AR165018.1 GI:16238344
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 80161)
AUTHORS Baltz,R.H., Crawford,K.P., Broughton,M.Christine., Madduri,K.,
Merlo,D.J., Turner,J.R., Treadway,P.J. and Waldron,C.
TITILE Biosynthetic genes for spinosyn insecticide production
JOURNAL Patent: US 6274350-A 1 14-AUG-2001;
FEATURES
Source
1..80161
Location/Qualifiers
1..80161
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 2,88e-290 Length: 80161
Score: 275.00 Matches: 275
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-069-353a-8 (1-275) x AR165018 (1-80161)

Qy 1 ValLeuProGlyValAlaProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThr 20
Db 20168 GTGTTGCCAGTGGCGGACCAACATCGCAGAGTTGGGAGATGTATGACTGGTCAAG 20227

Qy 21 ProLeuLeuAsnSerValAlaGlyGlyProCybAlaIleHisGlyTyrTrpGluAsn 40
Db 20228 CCGTTGCTGAACCTCGGTCGGGGGGGCGCCCTGCGCCATCCACACCGGCTACTGGAGAAC 20287

Qy 41 AspGlyAArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
Db 20288 GACGGGGGGGCTTCCTGCGAGCAGCGCGCGGACCGGCTGACCGACTTGTGCGCGAAGCG 20347

Qy 61 ThrValLeuAspGlyValAlaArgLeuLeuAspValGlyCybGlyThrGlyGlnProAla 80
Db 20348 ACCGTGCTGATGGCGGCGGCTTCGACTGCTCGATGTGGGGTGGCGTACCGGACACCAAGCG 20407

Qy 81 LeuArgValAlaAArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGln 100
Db 20408 CTCGGCGTCCGGCGGCAACCGGATCCAGATCACCGGCTACCGGCTGAGCA 20467

Qy 101 ValAlaIleAlaAlaAspCybAlaAArgGluArgGlyLeuSerHisArgValAspPheSer 120
Db 20468 GTGGCCATCGCCGCTGATGCGGACCGGAGCTAAGCCACCGGAGTGAATCTTCTCG 20527

Qy 121 CysValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
Db 20528 TGGGTGATGATCCATGTCCTCGCGGTAACCGGACATCTTTCAGCGCGCTGGGCATG 20587

Qy 141 GlnSerLeuLeuGlnMetSerGluProAspArgAlaIleArgGluIleLeuArgValLeu 160
Db 20588 CAGTCCGCTGTGAGATGTCGGAACCGGACCGGAGCTAAGCAATCTTTCAGATCACTC 20647
```

QY 161 LysProGlyGlyIleLeuGlyValThrGluValIleValArgGluIleGlyGlyMet 180
|||
Db 20648 AAACCCGGTGGACATCTCGCGCTGACCGAGTGTGTCAACAGAGACGGCGGCGGAGG 20707
QY 181 ProValSerGlyAspArgTrpProThrGlyLeuArgIleCysLeuAlaGluGlnLeu 200
20708 CCGGTGTCCGGGGACAGGTGGCCGACCGGCTTGAGATCTGCTGAGCAACTTCTG 20767
QY 201 GluSerLeuArgAlaIleGlyPheGluIleLeuAspTrpGluAspValSerSerArgThr 220
20768 GAATCCCTGCTGACACGGGGGTTCGAGATCTCGATTGGGAGGACGCTGTCTGAGGACC 20827
QY 221 ArgTyrPheMetProGlnPheAlaGluGlnLeuAlaIleHisGlnHisGlyIleAlaAsp 240
Db 20828 CGGTACTTCATGCGCGACGATTCGCGCAAGAGCTGCTGCGCACACGACGGAGTCCGAGAC 20887
QY 241 ArgTyrGlyProAlaValAlaIleGlyTyrAlaAlaAlaValCysAspTyrGluTyrAla 260
Db 20888 AGGTACGGGGCGGCTGTGCGCGGCTGGCGCGCGGCTGTGCGATTATGAGAAATATGCC 20947
QY 261 HisAspMetGlyTyrAlaIleLeuThrAlaArgLysProValGly 275
Db 20948 CACGACATGGGCTATGCGATTTCGACGGCGCGGACCGGTCGCGC 20992

RESULT 6
BD137649 80161 bp DNA linear PAT 18-SEP-2002
LOCUS Biosynthetic genes for spinosyn insecticide production.
DEFINITION BD137649
ACCESSION BD137649.1 GI:23232594
VERSION JP 2002505881-A/1.
KEYWORDS Saccharopolyspora spinosa
SOURCE Saccharopolyspora spinosa
ORGANISM Saccharopolyspora spinosa
Bacteria; Actinobacteria; Actinomycetales;
Pseudonocardiaceae; Pseudonocardiaceae; Saccharopolyspora.
REFERENCE 1 (bases 1 to 80161)
AUTHORS Baltz,R.H., Broughton,M.C., Crawford,K.P., Madduri,K., Merlo,D.J.,
Treadway,P.J., Turner,J.R. and Waldron,C.
TITLE Biosynthetic genes for spinosyn insecticide production
JOURNAL Patent: JP 2002505881-A 1 26-FEB-2002;
DOW AGROSCIENCES LLC
COMMENT OS Saccharopolyspora spinosa
PN JP 2002505881-A/1
PD 26-FEB-2002
PE 16-FEB-1999 JP 2000535754
PR 09-MAR-1998 US 09/036967
PI RICHARD H BALTZ,R M CHRISTINE BROUGHTON,KATHERYN P CRAWFORD, PI
KRISHNAMURTHY MADHURI,DONALD J MERLO,PATTI J TREADWAY,JAN R PI
TURNER,
PI CLIVE WALDRON
PC C12N15/09,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12P19/62,C12O1/
PC 68//
PC C07K14/195,C12N15/00,C12N5/00
CC Biosynthetic genes for spinosyn insecticide production FH
Key Location/Qualifiers
FT source 1..80161
FT location/Qualifiers
1..80161
1..80161
/organism="Saccharopolyspora spinosa"
/mol_type="genomic DNA"
/db_xref="taxon:60894"

ORIGIN
Alignment Scores:
Pred. No.: 2,88e-290 Length: 80161
Score: 275.00 Matches: 275
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
Db:

US-10-069-353a-8 (1-275) x BD137649 (1-80161)
QY 1 ValLeuProGlyGlyAlaProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThr 20
|||
Db 20168 GTGTGGCCAGGTGGGCGCACCAACATCGCAGCGGTGGGCAAGATGTATGACTGTGTCAGC 20227
QY 21 ProLeuLeuAsnSerValAlaGlyGlyProCysAlaIleHisGlyTyrTrpGluAsn 40
Db 20228 CCGTCTCTAATCTCGGTCCGGGCGGCGCCCTCGGCATTCACACCGGCTACTGGAGAAC 20287
QY 41 AspGlyArgAlaSerTrpGlnGlnAlaIleAspArgLeuThrAspLeuValAlaGluArg 60
Db 20288 GACGGGCGGGCTTCTTCTGGCAGCAGGCGCGCACCGGCTCACCGACTTGTCCGCAACGG 20347
QY 61 ThrValLeuAspGlyGlyValAlaArgLeuLeuAspValGlyCysGlyThrGlyGlnProAla 80
Db 20348 ACCGTGCTCGATGGGCGGCTTGTGACTGTCTGATGTGGGGGTGGGATCCGACACACACG 20407
QY 81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGln 100
Db 20408 CTGCGCGTGGCGCGCGCACACCGCATCCAGTACCGGCATCACCGTCAAGCCAGGTGCNA 20467
QY 101 ValAlaIleAlaIleAspCysAlaArgGluArgGlyLeuSerHisArgValAspSer 120
Db 20468 GTGGCCATCGCGCTGATTCGCGACGCGAACCGGACTTAAGCCACCGGGTGGACTTCTCG 20527
QY 121 CysValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaIleTrpAlaMet 140
Db 20528 TCGCTGATGACCATGTCTCTGCGTACCGGCAATGCTTTCGACGCGCGCTGGGCGCATG 20587
QY 141 GlnSerLeuLeuGluMetSerGluProAspArgAlaIleArgGluIleLeuArgValLeu 160
Db 20588 CAGTGGCTGTGAGATGTCCGAACCGGACCGGTGCATCCGGGAATCTTCGATCATCTC 20647
QY 161 LysProGlyGlyIleLeuGlyValThrGluValValLysArgGluAlaGlyGlyMet 180
Db 20648 AAACCCGGTGGATCTCGGCTCTACCGAGGTCTGCAACGAGACGGCGCGGCGGATG 20707
QY 181 ProValSerGlyAspArgTrpProThrGlyLeuArgIleCysLeuAlaGluGlnLeu 200
Db 20708 CCGGTGTCCGGGGACAGGTGGCCGACCGGCTTGAGATCTGCTGAGCAACTTCTG 20767
QY 201 GluSerLeuArgAlaIleGlyPheGluIleLeuAspTrpGluAspValSerSerArgThr 220
Db 20768 GAATCGCTCGGCGACGCGGGGTTCGAGATCTTCGATTGGAGAGCTGTCTCGAGGAC 20827
QY 221 ArgTyrPheMetProGlnPheAlaGluGlnLeuAlaIleHisGlnHisGlyIleAlaAsp 240
Db 20828 CCGTACTTCATGCCCGACATTCGCGCAAGAGCTCGCTGGCGCACACGACGGGATCCGGGAC 20887
QY 241 ArgTyrGlyProAlaValAlaIleGlyTyrAlaAlaAlaValCysAspTyrGluTyrAla 260
Db 20888 AGGTACGGGGCGGCTGTGCGCGGCTGGCGCGCGGCTGTGAGATTATGAGAAATATGCC 20947
QY 261 HisAspMetGlyTyrAlaIleLeuThrAlaArgLysProValGly 275
Db 20948 CACGACATGGGCTATGCGATTTCGACGGCGGGAACCGGTCGCGC 20992

RESULT 7
AR281866 80161 bp DNA linear PAT 10-APR-2003
LOCUS AR281866
DEFINITION Sequence 1 from patent US 6521406.
ACCESSION AR281866
VERSION AR281866.1 GI:29717767
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 80161)
AUTHORS Baltz,R.H., Broughton,M.C., Crawford,K.P., Madduri,K., Merlo,D.J.,
Treadway,P.J., Turner,J.R. and Waldron,C.
TITLE Spnd, a gene for spinosyn insecticide biosynthesis
JOURNAL Patent: US 6521406-A 1 18-FEB-2003;

FEATURES
source
Location/Qualifiers
1..80161
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 2,88e-290 Length: 80161
Score: 275.00 Matches: 275
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-069-353a-8 (1-275) x AR281866 (1-80161)

QY 1 ValLeuProGlyValAlaProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThr 20
DB 20168 GTGTGCGAGTGGCGGACCAACATCGACAGGTTGGCAGATGATGACTGTGCACG 20227
QY 21 ProLeuLeuAsnSerValAlaGlyValProCysAlaLeuHisGlyTyrTrpGluAsn 40
DB 20228 CCGTCTGTAACCTCGGCGGGCGGCCCTGCGCCATCCACACGCGTACTGGAGAAC 20287
QY 41 AspGlyValArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
DB 20288 GACGGCGGGCTTCTTGCGACGAGCGCGCCGACCGGCTCACCACTTGTCCCGAAGCG 20347
QY 61 ThrValLeuAspGlyValAlaArgLeuLeuAspValGlyCysGlyTyrTrpGlnProAla 80
DB 20348 ACCGTCTCGATGGCGGCGTTGACATGCTGATGTGGGGTGGCGTACCGGACCAACGCG 20407
QY 81 LeuArgValAlaArgAspAsnAlaGlnGlnLeuThrGlyTyrLeuThrValSerGlnValGln 100
DB 20408 CTGCGGCGTCCGGCCCAACAGCGGATCCAGATCACCGGCACTCCGTCACCGAGGTGCCA 20467
QY 101 ValAlaGlnAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValaAspPheSer 120
DB 20468 GTGGCCATCGCCCTGATTCGACCGGAAACCGGACTAAGCAACCGGGTGGACTTCTCG 20527
QY 121 CysValaAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
DB 20528 TGGGTGAGTGCATGTCCTCGCGGTACCGGACAACTGCTTCCACCGCGCTGGGCGCATG 20587
QY 141 GlnSerLeuLeuGlnMetSerGluProAspArgAlaAlaArgGlnLeuLeuValLeu 160
DB 20588 CAGTCCCTGTGGAGATGTCGGAACCGGACCGTGCATCCGGGAAATCCTTCCAGTACTC 20647
QY 161 LysProGlyValTyrLeuGlyValThrGluValValLysArgGluAlaGlyValGlyMet 180
DB 20648 AATCCCGGTGGCATCTCGCGGTCAACGAGGTGTCGAAGAGAGGCGGCGGCGGAGTGC 20707
QY 181 ProValSerGlyLysAspArgTrpProThrGlyLeuArgGlyLeuSerLeuAlaGlnLeuLeu 200
DB 20708 CCGGTGTCGGGAGACAGGTGGCGGACCGGCTTCGAGTCTGCTGCTGAGCAACTTCTG 20767
QY 201 GlnSerLeuArgAlaAlaGlyPheGlnLeuLeuAspTrpGluAspValSerSerArgThr 220
DB 20768 GAATCCCTGTGTCACGCGGGTTCGAGATCTCTGATTTGGGAGACAGCTGTCTGAGAGACC 20827
QY 221 ArgTyrPheMetProGlnPheAlaGlnGluLeuAlaAlaHisGlnHisGlyTyrLeuAlaAsp 240
DB 20828 CCGTACTTCAATGCGGAGTTGCGCGAAGACTCGCTGCGCACAGCAACGCGGATCGCGGAC 20887
QY 241 ArgTyrGlyProAlaValAlaGlyTyrAlaAlaAlaValaCysAspTyrGluLysTyrAla 260
DB 20888 AGGTACGGGCGGCTGTCGCGCGGCTGGCGCGCGCTGCGCATTAATGAGAAATATGCG 20947
QY 261 HisAspMetGlyTyrAlaAlaLeuThrAlaArgLysProValGly 275
DB 20948 CACGACATGGGCTATGAGATTGACGCGCGGAGCGGCTCGCG 20992

RESULT 8

AX600587/c
LOCUS AX600587 36538 bp DNA linear PAT 14-FEB-2003
DEFINITION Sequence 2 from Patent WO02079477.
ACCESSION AX600587
VERSION AX600587.1 GI:28400303
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
Dow Agrosciences LLC (US)
FEATURES
source
Location/Qualifiers
1..36538
/organism="Saccharopolyspora sp. NRRL 30141"
/mol_type="unassigned DNA"
/db_xref="taxon:221471"
/note="LM107129"

ORIGIN

Alignment Scores:
Pred. No.: 7,33e-29 Length: 36538
Score: 37.00 Matches: 37
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.45% Indels: 0
DB: Gaps: 0

US-10-069-353a-8 (1-275) x AX600587 (1-36538)

QY 7 ProThrSerGlnGlnValGlyGlnMetTyrAspLeuValThrProLeuLeuAsnSerVal 26
DB 920 CCACATCCACAGAGTCCGGCAGATGATGACTGTGACCCCGTGTGATTCGATC 861
QY 27 AlaGlyValProCysAlaAlaLeuHisGlyTyrTrpGluAsnAspGlyArg 43
DB 860 GCGGGCGGCGCCCTGCGCCATCCACACGCGTACTGGAGAACGACGCGCGG 810

RESULT 9
AR534398
LOCUS AR534398 738 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 13 from patent US 6733998.
ACCESSION AR534398
VERSION AR534398.1 GI:53924595
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
Dow Agrosciences LLC (US)
FEATURES
source
Location/Qualifiers
1..738
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 1.06 Length: 738
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.64% Indels: 0
DB: Gaps: 0

US-10-069-353a-8 (1-275) x AR534398 (1-738)

REFERENCE
1 (bases 1 to 738)
Thorson, J.S.
Micromonospora echinospora genes coding for biosynthesis of
calicheamicin and self-resistance thereto
Patent: US 6733998-A 13 11-May-2004;
Location/Qualifiers
1..738

Qy 68 ArgLeuLeuAspValGlyCysGlyThrGly 77
|||||
Db 121 CGGCTGCTCGACGTCGGCTGCCGACCGGC 150

RESULT 10
AX573660
LOCUS AX573660 738 bp DNA linear PAT 07-JAN-2003
DEFINITION Sequence 13 from Patent WO02079465.
ACCESSION AX573660
VERSION AX573660.1 GI:27551330
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Thorson, J.
TITLE Micromonospora echinospora genes encoding for biosynthesis of
JOURNAL calicheamicin and self-resistance thereto
MEMORIAL Patent: WO 02079465-A 13 10-OCT-2002;
SLOAN-KETTERING CANCER CENTER (US)
LOCATION/Qualifiers
1..738
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/note="Bacteria"
1..738
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD6016.1"
/db_xref="GI:27551331"
/translation="MGPEHALEYENAARGRGSWHDEADVADRIIPARPDARLLD
VCGGAGHLETPATRPVEGLAPAMALAHRLPGVRLHAGDMRTFDLGVTPDAV
TCLFTAVNLTGVAEMRAVAAMASHLAAGGLVLEPPMFPRLFDGYGGLVREG
RVAVRSRSTRGRVTRMERLNGDAIGIRFSQGLLMTREBYDAAPAAAGCES
ATVEGWLTRGFLVATRTGHALPTNV"

CDS
1..738
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD6016.1"
/db_xref="GI:27551331"
/translation="MGPEHALEYENAARGRGSWHDEADVADRIIPARPDARLLD
VCGGAGHLETPATRPVEGLAPAMALAHRLPGVRLHAGDMRTFDLGVTPDAV
TCLFTAVNLTGVAEMRAVAAMASHLAAGGLVLEPPMFPRLFDGYGGLVREG
RVAVRSRSTRGRVTRMERLNGDAIGIRFSQGLLMTREBYDAAPAAAGCES
ATVEGWLTRGFLVATRTGHALPTNV"

Alignment Scores:
Pred. No.: 1.06 Length: 738
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.64% Indels: 0
DB: Gaps: 0

US-10-069-353a-8 (1-275) x AX573660 (1-738)

Qy 68 ArgLeuLeuAspValGlyCysGlyThrGly 77
|||||
Db 121 CGGCTGCTCGACGTCGGCTGCCGACCGGC 150

RESULT 11
BT009270
LOCUS BT009270 788 bp mRNA linear PLN 20-JUN-2003
DEFINITION Triflicum aestivum clone wlx8.p0008.g5:fls, full insert mRNA
sequence.
ACCESSION BT009270
VERSION BT009270.1 GI:32128821
KEYWORDS
SOURCE Triflicum aestivum (bread wheat)
ORGANISM Triflicum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 788)
Tiney, S.V., Wolter, P., Powell, W., Dolan, M., Miao, G.-H.,
Caraher, N.R., Hanafey, M.K. and Hanney, C.F.
Direct Submission
Submitted (20-JUN-2003) Crop Genetics, E. I. Dupont de Nemours and
Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104,
USA
Location/Qualifiers

source 1..788
/organism="Triflicum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wlx8.p0008.g5:fls"

Alignment Scores:
Pred. No.: 1.12 Length: 788
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.64% Indels: 0
DB: Gaps: 0

US-10-069-353a-8 (1-275) x BT009270 (1-788)

Qy 148 GIUProApAgaAlIlleArgGluIleLeu 157
|||||
Db 541 GAGCCAGACCGCGCATTCGGAGATTG 570

RESULT 12
AY491379/c
LOCUS AY491379 1394 bp mRNA linear VRT 17-SEP-2004
DEFINITION Acanthopagrus schlegelii FTZ-F1 (f1b) mRNA, partial cds.
ACCESSION AY491379
VERSION AY491379.1 GI:45686254
KEYWORDS
SOURCE Acanthopagrus schlegelii (black sea bream)
ORGANISM Acanthopagrus schlegelii
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;
Percoidae; Sparidae; Acanthopagrus.
1 (bases 1 to 1394)
Lin, X., Liang, B. and Zhang, S.
Sequence and expression of cytochrome P450 aromatase and FTZ-F1
genes in the protandrous black porgy (Acanthopagrus schlegelii)
Gen. Comp. Endocrinol. 138 (3), 247-254 (2004)

JOURNAL
PUBMED
15364207
REFERENCE 2 (bases 1 to 1394)
AUTHORS Xusheng, L., Bing, L. and Shuyi, Z.
TITLE Direct Submission
JOURNAL Submitted (28-NOV-2003) Institution of Zoology, Chinese Academy of
Sciences, 25 Beisihuanxilu, Haidian, Beijing 100080, China

FEATURES
source
1..1394
/organism="Acanthopagrus schlegelii"
/mol_type="mRNA"
/db_xref="taxon:72011"
1..>1394
/gene="f1b"
/note="afflb"
1..>1394
/gene="f1b"
/note="Orphan nuclear receptor"
/codon_start=1
/product="FTZ-F1"
/protein_id="AA575792.1"
/db_xref="GI:45686255"
/translation="MLGDKAQGVTLKWEYTYDDLEELCYVCGDKVSGHYGLITCE
SKGFFQRTVONNRKYTCENQECKIDTQRKCPFCRFOKLSVGMLEAVRADRM
GGRNFGMYNDRDLAKQKALIRSNQFKLSVAPPASLQTDYGFGLHSLPTI
SKLLSPRSSITPTDYSALYGPSPGMAQSHVPLTQYOTTFPGRAIKAECPDY
TSSPSRLTGYPYPMYSPASPOPSLPPIVLELARCDSGGARLDKPSFSLMCMAD
QTLRSIVEMARSCIFFKELRVGDQKLIHNCWSBLVDHIIFROYOHCQESILVTC
QAEVLSFTLSQSEATLSLVORGOELAVRLALQVDRVACLKFLLFNPVKLLED
QAFEGVQEOVNGALLLEYLTSTYPOFOKFSQLVRLPRLSLSTQARDYLCYMLSG
EVPNNLLIEMLHA"

ORIGIN
Alignment Scores:
Pred. No.: 1.88 Length: 1394

Score: 10.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.64%
Matches: 10
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-069-353A-8 (1-275) x AY491379 (1-1394)

QY 244 ProAlaValAlaGlyTrrpAlaAlaVal 253
DB 742 CACGGCGTGCAGCGCTGCGCGCGG 713

RESULT 13
A60304
LOCUS A60304 6085 bp DNA linear PAT 06-MAR-1998
DEFINITION Sequence 4 from Patent WO9708323.
ACCESSION A60304
VERSION A60304.1 GI:3715279
KEYWORDS
SOURCE unidentified
ORGANISM unidentified

REFERENCE 1
AUTHORS Schupp,T., Engel,N., Bietenhader,J., Toupet,C. and Pospiech,A.
TITLE STAUROSFORIN BIOSYNTHESIS GENE CLUSTERS
JOURNAL Patent: WO 9708323-A 4 06-MAR-1997;
CIBA GEIGY AG (CH)

FEATURES
source Location/Qualifiers
1..6085
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

misc_RNA 378..1665
/function="ORF"
misc_RNA 1747..2553
/function="ORF"
misc_RNA 2593..4011
/function="ORF"
misc_RNA 4013..4999
/function="ORF"
misc_RNA 5071..6085
/function="ORF"

ORIGIN
Alignment Scores:
Pred. No.: 7.12 Length: 6085
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatch: 0
Query Match: 3.64% Indels: 0
Gaps: 0

US-10-069-353A-8 (1-275) x A60304 (1-6085)

QY 51 AspArgLeuThrAspLeuValAlaGluArg 60
DB 1894 GACCGGCTCACCAATCTCGTGGCGAGCGC 1923

RESULT 14
AR144763
LOCUS AR144763 6085 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 4 from patent US 6210935.
ACCESSION AR144763
VERSION AR144763.1 GI:15106630
KEYWORDS
SOURCE Unknown.
ORGANISM Unidentified.

REFERENCE 1 (bases 1 to 6085)
AUTHORS Schupp,T., Engel,N., Bietenhader,J., Toupet,C. and Pospiech,A.
TITLE Stauroporin biosynthesis gene clusters
JOURNAL Patent: US 6210935-A 4 03-APR-2001;
FEATURES Location/Qualifiers

source 1..6085
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 7.12 Length: 6085
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatch: 0
Query Match: 3.64% Indels: 0
Gaps: 0

US-10-069-353A-8 (1-275) x AR144763 (1-6085)

QY 51 AspArgLeuThrAspLeuValAlaGluArg 60
DB 1894 GACCGGCTCACCAATCTCGTGGCGAGCGC 1923

RESULT 15
AE005764
LOCUS AE005764 10802 bp DNA linear BCT 12-JUN-2002
DEFINITION Caulobacter crescentus CB15 section 90 of 359 of the complete genome.
ACCESSION AE005764
VERSION AE005764.1 GI:13422130
KEYWORDS
SOURCE Caulobacter crescentus CB15
ORGANISM Caulobacter crescentus CB15

REFERENCE 1
AUTHORS Nierman,W.C., Feldblyum,T.V., Laub,M.T., Paulsen,I.T., Nelson,K.E., Eisen,J., Heidelberg,J.F., Alley,M.R., Ohta,N., Maddock,J.R., Potocka,I., Nelson,W.C., Newton,A., Stephens,C., Phadke,N.D., Ely,B., Debroy,R.T., Dodson,R.J., Durkin,A.S., Gwinn,M.L., Haft,D.H., Kolonay,J.F., Smit,J., Craven,M.B., Khouri,H., Shetty,J., Berry,K., Uterback,T., Tran,K., Wolf,A., Vamathevan,J., Ermolaeva,M., White,O., Salzberg,S.L., Venter,J.C., Shapiro,L. and Fraser,C.M.
TITLE Complete genome sequence of Caulobacter crescentus
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (7), 4136-4141 (2001)
MEDLINE 21173698
PubMed 11259647

REFERENCE 2 (bases 1 to 10802)
AUTHORS Nierman,W.C., Feldblyum,T.V., Paulsen,I.T., Nelson,K.E., Eisen,J., Heidelberg,J.F., Alley,M.R., Ohta,N., Maddock,J.R., Potocka,I., Nelson,W.C., Newton,A., Stephens,C., Phadke,N.D., Ely,B., Laub,M.T., Debroy,R.T., Dodson,R.J., Durkin,A.S., Gwinn,M.L., Haft,D.H., Kolonay,J.F., Smit,J., Craven,M., Khouri,H., Shetty,J., Berry,K., Uterback,T., Tran,K., Wolf,A., Vamathevan,J., Ermolaeva,M., White,O., Salzberg,S.L., Shapiro,L., Venter,J.C. and Fraser,C.M.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2001) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA
FEATURES Location/Qualifiers
1..10802
/organism="Caulobacter crescentus CB15"
/mol_type="genomic DNA"
/strain="CB15"
/db_xref="taxon:190650"
complement(86..262)
/gene="CC0870"
complement(86..262)
/gene="CC0870"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK2855.1"
/db_xref="GI:13422131"
/translation="MAFKRAALGALMLVAGAAQAAMVTPAVGQPLARFDKLRPAS


```
gene      PAPDGSVWVKTLIM"
CDS        complement (341. .925)
           /gene="CC0871"
           /complement (341. .925)
           /note="identified by Glimmer2; putative"
           /codon_start=1
           /transl_table=11
           /product="hypothetical protein"
           /protein_id="AAK22856.1"
           /db_xref="GI:13422132"
           /translation="MLIGRAASSLPPEPLMNI.PMPLILTYKDDIDLADIALVAAGVAG
           ETRAQQTISESELGGQRTDEASLRADLAKAGSTLLTFRDAMWKEPFCAMITPAD
           DWYVMYITVPTROQAGMGRAVLLETCEAYARAGKGVARMVTIVISVRDALIMYBERG
           YLUTGETOPFPYGDDRFGLPQRDDLAFFVMEKGL"
           964. .1761
           /gene="CC0872"
           964. .1761
           /gene="CC0872"
           /note="identified by Glimmer2; putative"
           /codon_start=1
           /transl_table=11
           /product="hypothetical protein"
           /protein_id="AAK22857.1"
           /db_xref="GI:13422133"
           /translation="MPDLBSRPVMVRAQARIVSLYRHPYKGFPTPELERAADIAGAFC
           PCDRLYAVEDSGSDPDAAPAHISKKFTVLAKIPAVARATAYDEKSGVLTAAEGR
           PPAGDLROASGRAGFEALDLGLDAGCPKVI.EGPGARFPMDSKGVYITVNLAS
           VADLSAKVGRDLDPLRFRAMLTYEGPANAENDLTGQTLISGAKRAVTKPIVRCAAT
           HVDPTSAERDIDLVKALFDHGHMFCGIYLVNTEGDAVREGDVAAAG"
           complement (1762. .2052)
           /gene="CC0873"
           complement (1762. .2052)
           /gene="CC0873"
           /note="identified by Glimmer2; putative"
           /codon_start=1
           /transl_table=11
           /product="conserved hypothetical protein"
           /protein_id="AAK22858.1"
           /db_xref="GI:13422134"
           /translation="MKPYRLSRRAKADLDITTYSBORWGVQADAVARELOATTEMI
           AHPGWDQDENLRAGYRCASGHVFVAVGVEIIRVLHQSNNARHLG"
           complement (2049. .2315)
           /gene="CC0874"
           complement (2049. .2315)
           /gene="CC0874"
           /note="identified by Glimmer2; putative"
           /codon_start=1
           /transl_table=11
           /product="conserved hypothetical protein"
           /protein_id="AAK22859.1"
           /db_xref="GI:13422135"
           /translation="MAKSKTSTVVLGDHFOAFTDSQVADGRYSASBEVIRAGLRLLSEN
           EKLAALRALALEGESGFIEDFDFALEERSASAPGFHEE"
           2305. .3258
           /gene="CC0875"
           2305. .3258
           /gene="CC0875"
           /note="identified by match to protein family HMM"
           /codon_start=1
           /transl_table=11
           /product="hemk family protein"
           /protein_id="AAK22860.1"
           /db_xref="GI:13422136"
           /translation="MLAMADLYQNLVPHSISPPGRRRAKAPMTLTLYKAWTAADRL
           KQAGIDPSIDARLML.EVAVAGVTRTEIVDPYRELASAOIATLNDYLERARREPVSH
           IIRGKFWKILIQVNNKNTLPPEPTEVIVDEVLKAFPEBMATSM.LDGVSGSTIILAV
           IAEPRAAKGLIDIASSEALVARENNANLIDNTRALLHGDWTTGSGSPPLVAVSNP
           PYTPTEVDTLBEVRIHEPRLADGGPGLGAAYREIAPBEIIRVAKPGGLFAVEIGID
           OSQAVALEPRAAGATEVTRVKDLSHDRVVLGVKAPLESPA"
           3517. .4629
           /gene="CC0876"
           3517. .4629
```

```
gene      /gene="CC0876"
CDS        /note="identified by Glimmer2; putative"
           /codon_start=1
           /transl_table=11
           /product="hypothetical protein"
           /protein_id="AAK22861.1"
           /db_xref="GI:13422137"
           /translation="WRDFKMKRORGRUNRGNGGKPOOHANRAFDNSNGPEGVKVRG
           AAOVYKEYOOLARPASSGDRVLAENYLOHAHEHFRVLRAIOPVRPAVDIVGKVVA
           AYIEDFEAPPEPEPVEVQPAAGQAGEGNGENPRDRPENRBRNDRDRNENRD
           KDRNDDRNKRGDSRRDRKDRDDPRDRPRDPRDRPRDRPRDRPRDRPRDRPRDR
           DFRDRDRPRDRPRDRPAAGEQPGEGEFRSRRRDRNRDRDRDWOPEGRSDP
           MAVIEDPASP.LTAEAPAGEGSPMLRGDGAASHAPAFYGRARERAPAEAPAEAA
           PASDGECEVPRRRARPSFEAPAEAGABEGSDGC"
           4797. .5552
           /gene="CC0877"
           4797. .5552
           /gene="CC0877"
           /note="identified by Glimmer2; putative"
           /codon_start=1
           /transl_table=11
           /product="conserved hypothetical protein"
           /protein_id="AAK22862.1"
           /db_xref="GI:13422138"
           /translation="WTLPELILPLILAAACALGGCAVNPDIQRDQLIIVNDNLARDG
           EKAMADALRTSHVSKKEPKNARVOVAGRVVAAAGLGRPWEVAFLEAPNAFVLP
           GVGWATVALLSLVNDQDLAVIGHGAGVVARHRAAEVSOETTSKVILGAKAAGG
           TFEGLLDHDEHDAIRYVLLIPSKSKOLEMDKLGVDVMQAGIRPREAVVLNRMQA
           LDGKSGGDLDS.THPSDAVRIQELERYIASKGM"
           5651. .8230
           /gene="CC0878"
           5651. .8230
           /gene="CC0878"
           /note="identified by match to protein family HMM"
           /codon_start=1
           /transl_table=11
           /product="ATP-dependent Clp protease, ATP-binding subunit
           ClpB"
           /protein_id="AAK22863.1"
           /db_xref="GI:13422139"
           /translation="WNIDLVSDRAKQAVOS.AOSLALAHQOAPAEHILKLYLEBKDG
           LSRALIQAGSRPDQLDGVVETLLAKTPRVDAAGQOLYMKDPTARVFEAKSKAAAG
           DAFVTERLLI.AIAKEGGEAAKLFEKAGVSAQSLETANANRKRGTADSAABAEVYA
           LKRYVARDLTAARDOCKLDPVIGRDEIRRTIOVLSRTKXNPVLIGEGVGKTAIVEG
           LAIRTVNGDVPESILDKKLLSLDMSGLIAGKVGPEPERKAVIGEVTAEGSTIIF
           IDEKATVYAGKGDGANDASNLKLRALARGELHCYGAATTDEYRKHYEKDALARRQ
           PFVSEPTVEDTVSILRGLEKEVEHGVRLSDSAIVAAATLSNRYIADRFLPDPAID
           LVDEASRVRMOIDSKEPELDEIDRLVQLTERLASKETDASKORLENI.EVEIID
           LQFSDDEMTARMAKEKEKVGGAQAREALDLRADLANAQAQDPAFARAQIOYGEIIPA
           LERRIAEAAGDTOLATPEVVDAEOIAVVRWMTGVPEKMLEGREGKTLAMEDELARG
           RVGQDEALEAVSADVAPRAGLQDPSKISGPIFGPTGVGKTEILSLAEPLFADP
           AAITMDSEYMEKHSVSR.LIGAPGYGVDEGALTAIRRPPIQVILFPEIEKAP
           DVFNVLLOVLDLDRGTDSQKRTVPFRNTLLITWSNLGEAVLSAQDSQDRAVALAPVW
           NTVRGHPFERPLNRIIDELILFRKRLSNHMGIOIVIQLRVSKLLADGRMALALAEAL
           NMLADKGYDPVYGARPLKRVIOKELVPIAKKLAGEIEDGVIAGVTQQLSPGKA
           VLOQ"
           8231. .8782
           /gene="CC0879"
           8231. .8782
           /gene="CC0879"
           /note="identified by Glimmer2; putative"
           /codon_start=1
           /transl_table=11
           /product="hypothetical protein"
           /protein_id="AAK22864.1"
           /db_xref="GI:13422140"
           /translation="WTSLSPPCGGMPRRKPGSGELKTI.SAATGLSAGRGFGRSDP
           SSVAFSDPFSKGRGRASRP.PALASPRMAKTKLILPEPIEVGDGVYLRHLRPLP
           WSYDPSBHGKVIAARHKGOREALIEPFPKRFKFSAPLAKPSOIPRIYAKAIASA
           IEAGWEPUSRKGAVVIYDEEGA"
           complement (8792. .9070)
           /gene="CC0880"
           complement (8792. .9070)
```

```
/gene="CC0880"  
/note="identified by Glimmer2; putative"
```

/note="identified by Glimmer2; putative"

Alignment Scores:

Pred. No.:	12	Length:	10802
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.64	Indels:	0
DB:	1	Gaps:	0

Score:	10.00	Matches:	10
--------	-------	----------	----

Percent Similarity:	100.00%	Conservative:	0
Best local Similarity:	100.00%	Mismatches:	0

```

2000 .....: 200:000 .....
Query Match: 3.64% Indels: 0

```

DB:	1	Gaps:	0
-----	---	-------	---

US-10-069-353A-8 (1-275) x AE005764 (1-10802)

155 Clujul de Sus, 15 iunie 1964

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 29, 2005, 20:03:44 ; Search time 611 Seconds
(without alignments)
2664.368 Million cell updates/sec

Title: US-10-069-353a-8

Perfect score: 275
Sequence: 1 VILGGAGPTSGQVGMVLT.....YKVAHDMGVALITARKPVG 275

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 segs, 2959870667 residues

Word size: 1

Total number of hits satisfying chosen parameters: 8776198

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/uspto.spool/US10069353.rnuc 26082005 172151 15042/app.query.fasta_1.455
-DB=N_Geneseq_16Dec04 -OPMT=fastap -SUFFIX=ring -MIMATCH=0.1 -LOOFCI=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human0.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=200000000
-USR=US10069353 @CGN 1 1 470 @runat 26082005 172151 15042 -NCPU=6 -ICPU=3
-NO MAMP -IARG=OLIBRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DET TIMEOUT=120 -MAMP TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6
-Fgapext=7 -Ygapop=60 -Ygapext=60 -Delop=6 -Delext=7

Database :

N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001as:*
6: geneseqn2002as:*
7: geneseqn2002as:*
8: geneseqn2003as:*
9: geneseqn2003as:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	275	100.0	828	4	AAF88318
2	275	100.0	45624	4	AAF88315
3	275	100.0	50000	4	AAF88312
4	275	100.0	80161	2	AAZ21501
5	37	13.5	36538	10	ABV75558

Result No.	Score	Query Match	Length	DB ID	Description
6	351	9	ACU22474		Ac122474 DNA clone
7	397	9	ACU22489		Ac122489 DNA clone
8	406	9	ACU22477		Ac122477 DNA clone
9	439	9	ACU22483		Ac122483 DNA clone
10	552	9	ACU22491		Ac122491 DNA clone
11	572	9	ACU22487		Ac122487 DNA clone
12	607	9	ACU22468		Ac122468 DNA clone
13	617	9	ACU22454		Ac122454 DNA clone
14	632	9	ACU22486		Ac122486 DNA clone
15	633	9	ACU22481		Ac122481 DNA clone
16	634	9	ACU22484		Ac122484 DNA clone
17	642	9	ACU22476		Ac122476 DNA clone
18	651	9	ACU22473		Ac122473 DNA clone
19	675	9	ACU22488		Ac122488 DNA clone
20	685	9	ACU22475		Ac122475 DNA clone
21	700	9	ACU22482		Ac122482 DNA clone
22	703	9	ACU22478		Ac122478 DNA clone
23	703	9	ACU22458		Ac122458 DNA clone
24	705	9	ACU22490		Ac122490 DNA clone
25	708	9	ACU22485		Ac122485 DNA clone
26	719	9	ACU22472		Ac122472 DNA clone
27	721	9	ACU22479		Ac122479 DNA clone
28	726	9	ACU22480		Ac122480 DNA clone
29	738	8	ABX56032		Abx56032 M. echino
30	6085	2	AAT70153		Aat70153 S. longisp
31	759	8	ACF73542		Acf73542 Staphyloc
32	11466	2	AAV74755		AAV74755 Staphyloc
33	47745	11	ACN44160		Acn44160 Mouse gen
34	64492	12	ADK16023		Adk16023 Streptomy
35	87	8	ABZ09681		Abz09681 Human cell
36	87	10	ABN71244		Abn71244 Tumour su
37	213	6	ABN71244		Abn71244 Streptococ
38	270	6	ABE65692		AbE65692 Lung canc
39	270	6	ABK64644		AbK64644 Human ben
40	270	6	ABV77957		Abv77957 Hypoxia-r
41	349	5	ABV07585		Abv07585 Human pro
42	382	10	ADG38142		Adg38142 Aspergill
43	405	5	ABV77517		Abv77517 Human pro
44	414	11	ABD01730		Abd01730 Pseudomon
45	441	10	ABX07019		Abx07019 S. pneumo

ALIGNMENTS

RESULT 1
AAF88318
ID AAF88318 standard; DNA; 828 BP.
XX
AC AAF88318;
XX
DT 28-AUG-2001 (first entry)
XX
XX S. spinoza DNA fragment encoding ORF1, SEQ ID 7.
DE
XX Porosamine; trimethylchitosane; polyketide synthase; biosynthesis;
KW spinosyn; polyketide aglycone; transgenic plant; insect resistance;
KW metrolide; insecticide; O-methyltransferase; ds.
XX
XX Saccharopolyspora spinosa.
OS
XX
XX DE19957268-A1.
PN
XX 08-MAR-2001.
PD
XX 29-NOV-1999; 99DE-01057268.
PF
XX 27-AUG-1999; 99DE-01040596.
PR
XX (FARB) BAYER AG.
PA
XX Ebertz G, Moehrle V, Froede R, Velten R, Salas JA;
PI
XX WPI; 2001-267102/28.
DR

DR P-PSDB: AAB70948.
 XX
 PT New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for recombinant production of insecticidal spinosyns and their derivatives.
 XX
 PS Claim 7a; Page 102-104; 354pp; German.
 XX
 CC This invention describes a novel method nucleic acid (I) and its encoded polypeptide (II) containing at least one region that encodes an enzymatic activity involved in biosynthesis of spinosyns. (I) are used (i) to identify, inactivate or modulate genes involved in the biosynthesis of (ii); (iii) to generate a library of polypeptide synthases; (iii) for adding forosamine or trimethylrhamsome to a spinosyn or polypeptide aglycone; and (iv) for recombinant production of the corresponding enzymes, which are used for production of (II), their precursors or derivatives, including production of transgenic plants that express (II) and thus have increased resistance to insects. (i) are also useful as markers for sequencing the Saccharopolyspora spinosa genome. (ii) are macroclides with insecticidal, but not antibacterial, activity, and can also be used to raise specific antibodies, useful for identifying expression clones in a gene bank. Cells transformed with (I) may produce (II) at significantly increased levels or produce new derivatives of (II). This sequence, ORF 1, encodes an S. spinosa O-methyltransferase protein

Alignment Scores:
 Pred. No.: 3,88e-265 Length: 828
 Score: 275.00 Matches: 275
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-10-069-353a-8 (1-275) x AAF88318 (1-828)

QY 1 ValLeuProGlyValAlaProThrSerGlnGlnValAlaGlyGlnMetTyrAspLeuValThr 20
 Db 1 GTGTGGCAAGGTGGCGCAACCAATCGACAGGTGGAGTATGACCTGGTACG 60
 QY 21 ProLeuLeuAsnSerValAlaGlyGlyProCysAlaIleHisIshGlyTyrTrpGluAsn 40
 Db 61 CCGTGGCTGAACCTGGTGGGGGGGGCCCTGGCCATCCACACCGGCTACTGGGAAAC 120
 QY 41 AspGlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGluArg 60
 Db 121 GACGGGGGGGCTTCTCTGGCAGACGGCCGCGACCGGCTCAACCGCTTGGCCGAAACGG 180
 QY 61 ThrValLeuAspGlyValAlaArgLeuLeuAspValAlaGlyCysGlyThrGlyGlnProAla 80
 Db 181 ACCGTGCTCAATGGCGGCGTTCGACTCGATGTGGGGTGGGATCCGACCAACGACGG 240
 QY 81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGln 100
 Db 241 CTGGGGCTCGGGCGCGCAACGCGATCCAGATCACGGGCAATCCGTACGACGATGCA 300
 QY 101 ValAlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120
 Db 301 GTGGCCCTCGCCGCTGATTGGCGACGGGAACGCGAATTAAGCCACCGGGTGGACTTCTCG 360
 QY 121 CysValAlaAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTPAlaMet 140
 Db 361 TGGCTCATGTCATGCTCCCTGCGCTACCGGCAATCTTTCGACGGCGGCTGGCCATG 420
 QY 141 GlnSerLeuLeuGluMetSerGluProAspArgAlaIleArgGluIleLeuArgValLeu 160
 Db 421 CAGTCCGTGTGGAGATGTCGGAACCGGACCGTCCATCCGGGAAATCTTCAGATCTC 480
 QY 161 LysProGlyGlyIleLeuGlyValAlaThrGluValAlaLysArgGluAlaGlyGlyMet 180
 Db 481 AAACCCCGTGGATCTCTCGGGGTTCACGAGAGTGTCTCAAAAGAGACGGCGGGGATG 540

QY 181 ProValSerGlyAspArgTrpProThrGlyLeuArgIleCysLeuAlaGluGlnLeu 200
 Db 541 CCGGTGTCGGGAGACAGTGGCCGACCGGCTTGGATCTGCTGGCTGACCACTTCTG 600
 QY 201 GluSerLeuArgAlaAlaGlyPheGluIleLeuAspTrpGluAspValSerSerArgThr 220
 Db 601 GAATCCGCTGCTGCACAGCGGGGTTCCGATCTCGATTGGAGAGACCTGTCTGAGAGACC 660
 QY 221 ArgTyrPheMetProGlnPheAlaGluGluLeuAlaAlaHisGlnHisGlyIleAlaAsp 240
 Db 661 CGGTACTTCAATGCGCAGATTGCGCGAAAGCTGCTGCGCACACGACGCGGATCGCGGAC 720
 QY 241 ArgTyrGlyProAlaValAlaGlyTyrAlaAlaIleValCysAspTyrGlyLysTyrAla 260
 Db 721 AGGTACGGGCGGCGTGTGCGCGGCTGGCCGCGCGGCTGCGATTATGAAATATGCC 780
 QY 261 HisAspMetClyTyrAlaIleLeuThrAlaArgLysProValGly 275
 Db 781 CAGGACATGGGCTATGCAATTCGACGGCGGGAACCGGTGGC 825

RESULT 2
 AAF88315/c
 ID AAF88315 standard; DNA; 45624 BP.
 XX
 AC AAF88315;
 XX
 DT 28-AUG-2001 (first entry)
 XX
 DE S. spinosa DNA fragment SEQ ID 4.
 XX
 KW Forosamine; trimethylrhamsome; polypeptide synthase; biosynthesis;
 KW spinosyn; polypeptide aglycone; transgenic plant; insect resistance;
 KW macroclide; insecticidal; ds.
 XX
 OS Saccharopolyspora spinosa.
 XX
 PN DE19957268-A1.
 XX
 PD 08-MAR-2001.
 XX
 PF 29-NOV-1999; 99DE-01057268.
 XX
 PR 27-AUG-1999; 99DE-01040596.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Eberz G, Moehle V, Froede R, Velten R, Salas JA;
 XX
 DR WPI; 2001-267102/28.
 XX
 PT New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for recombinant production of insecticidal spinosyns and their derivatives.
 XX
 PS Claim 7; Page 58-74; 354pp; German.
 XX
 CC This invention describes a novel method nucleic acid (I) and its encoded polypeptide (II) containing at least one region that encodes an enzymatic activity involved in biosynthesis of spinosyns. (I) are used (i) to identify, inactivate or modulate genes involved in the biosynthesis of (ii); (iii) to generate a library of polypeptide synthases; (iii) for adding forosamine or trimethylrhamsome to a spinosyn or polypeptide aglycone; and (iv) for recombinant production of the corresponding enzymes, which are used for production of (II), their precursors or derivatives, including production of transgenic plants that express (II) and thus have increased resistance to insects. (i) are also useful as markers for sequencing the Saccharopolyspora spinosa genome. (ii) are macroclides with insecticidal, but not antibacterial, activity, and can also be used to raise specific antibodies, useful for identifying expression clones in a gene bank. Cells transformed with (I) may produce (II) at significantly increased levels or produce new derivatives of (II). This sequence represents a genomic DNA fragment of the S. spinosa genome which contains the coding regions for proteins involved in forosamine and trimethylrhamsome biosynthesis

XX Sequence 45624 BP; 7933 A; 14913 C; 14940 G; 7838 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,74e-263	Length:	45624
Score:	275.00	Matches:	275
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-10-069-353A-8 (1-275) x AAF88315 (1-45624)

```

QY 1 valleuproglyglyalaprothserglnlnvalglylmettyrAspleuValThr 20
DB 828 GTGTTGCCAGTGGCCGACCAACATCGCAGAGTTGGGACAGTGTATGACCTGGTCAAG 769
QY 21 ProlenleuanserValalaglyglyProcybalailehihiaglytyrTTPguuAn 40
DB 768 CCGTGTCTGAATCGGTCCGGGGGGCCCTCGCATCCACCGGCTACTGGAGAAC 709
QY 41 AapglyArgAlaserTTPglnlnlnalalAAspargleuThrAspleuValalagluArg 60
DB 708 GACGGGGGGGCTTCTCGGACGAGCGCCGCGACCGGCTCACGACCTTGTCCGGAACGG 649
QY 61 ThrValleuAspGlyglyValArgleuAspValglyCyaglytyrTglnProAla 80
DB 648 ACCGTGCTCGATGGCGCGCTTCCACTGCTGATGGGGGTGCGGTACCGGACCAACGACGG 589
QY 81 leuArgValAlaArgaphenAlaleglnllethgyltlethralserglnlnValgln 100
DB 588 CTGGCGGTGGCGGCGACCAACCGGATCCAGATCCAGGACCGGACCGGACGAGTCA 529
QY 101 ValAlaileAlaAlaAspCysAlaArggluArgglyleuSerhisArgValAspPheSer 120
DB 528 GTGGCATTGCCCTGATATGGGACCGGACCGGACCGGACCGGACCGGACCGGACCGG 469
QY 121 CybValAspAlaWeSerleuProtyrProAspAsnAlaPheAspAlaAlaTTPAlaMet 140
DB 468 TGGGTGGATGGCCATGTCCTCGCGGACCGGACCGGACCGGACCGGACCGGACCGG 409
QY 141 GlnSerleuGlnleuSerGlnleuProAspArgAlaileArgglnlleuArgValleu 160
DB 408 CACTGCTGTGGAGATGTCGAGACCGGACCGGACCGGACCGGACCGGACCGGACCGG 349
QY 161 LysPProGlyglylleuGlyValAlaThrGluValAlaValysArggluAlaglyglylmet 180
DB 348 AAACCGGATGGCATCTCTCGGCGTCAACGAGTCTCTCAACGAGAACCGGCGGCGGATG 289
QY 181 ProValSerGlyAspArgTTPProthrglyleuArgglyleuAlaglnlnleu 200
DB 288 CCGGTGTCCGGGACAGGTGGCGACCGGCGCTTCCGATCGCTGCGTGGCAACTTCTG 229
QY 201 GlnSerleuArgAlaAlaglyPheGlnlleuAspTTPglnlnAspValSerArgTThr 220
DB 228 GAATCGCTGGGTGACAGGGGGTTCGAGATCTCGATGGGAGAGACGCTGTGAGAGACC 169
QY 221 ArgTy+PheMetProGlnPheAlagluGlnleuAlaAlaHisglnhisgylleAlaAsp 240
DB 168 CCGTACTTCAATGCCGAGTTCGCGGACGAGTCTCGCTCGGACCGGACCGGACCGG 109
QY 241 ArgTyglyProAlaValAlaAlaglyTTPAlaAlaAlaValCysAspTyrglylytyrAla 260
DB 108 AGGATCGGGCGCGGTGTGCGCGGCTGGGCGCGGCTGTGCAATTTGAGAAATATGCC 49
QY 261 HisAspMetGlytyrAlaileuThrAlaArglyserProValgly 275
DB 48 CACGACATGGGCTATGCGATTTCGACGGCGGAGACCGGATCGGC 4

```

RESULT 3
AAF88312
ID AAF88312 standard; DNA; 50000 BP.
XX

```

AC AAF88312;
XX 28-AUG-2001 (first entry)
XX S. spinosa DNA fragment SEQ ID 1.
XX Forosamine; trimethylrhinmose; polyketide synthase; biosynthesis;
XX spinosyn; polyketide aglycone; transgenic plant; insect resistance;
XX macrolide; insecticidal; de.
XX Saccharopolyspora spinosa.
XX DE19957268-A1.
XX 08-MAR-2001.
XX 29-NOV-1999; 99DE-01057268.
XX 27-AUG-1999; 99DE-01040596.
XX (FARB ) BAYER AG.
XX Eberz G, Moehrle V, Froede R, Velten R, Salas JA;
XX WPI; 2001-267102/28.
XX Claim 7, Page 14-31; 354p; German.
XX This invention describes a novel method nucleic acid (I) and its encoded
XX polypeptide (II) containing at least one region that encodes an enzymatic
XX activity involved in biosynthesis of spinosyns. (I) are used (i) to
XX identify, inactivate or modulate genes involved in the biosynthesis of
XX (II); (ii) to generate a library of polyketide synthases; (iii) for
XX adding forosamine or trimethylrhinmose to a spinosyn or polyketide
XX aglycone; and (iv) for recombinant production of the corresponding
XX enzymes, which are used for production of (II), their precursors or
XX derivatives, including production of transgenic plants that express (II)
XX and thus have increased resistance to insects. (I) are also useful as
XX markers for sequencing of the Saccharopolyspora spinosa genome. (II) are
XX also used to raise specific antibodies, useful for identifying
XX expression clones in a gene bank. Cells transformed with (II) may produce
XX (II) at significantly increased levels or produce new derivatives of
XX (II). This sequence represents a genomic DNA fragment of the S. spinosa
XX genome which contains the coding regions for proteins involved in
XX forosamine, trimethylrhinmose and polyketide synthase biosynthesis
XX
XX SQ Sequence 50000 BP; 8484 A; 16290 C; 16535 G; 8691 T; 0 U; 0 Other;

```

Alignment Scores:

Pred. No.:	1,89e-263	Length:	50000
Score:	275.00	Matches:	275
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-10-069-353A-8 (1-275) x AAF88312 (1-50000)

```

QY 1 valleuproglyglyalaprothserglnlnvalglylmettyrAspleuValThr 20
DB 44797 GTGTTGCCAGTGGCCGACCAACATCGCAGAGTTGGGACAGTGTATGACCTGGTCAAG 44856
QY 21 ProlenleuanserValalaglyglyProcybalailehihiaglytyrTTPguuAn 40
DB 44857 CCGTGTCTGAATCGGTCCGGGGGGCCCTCGCATCCACCGGCTACTGGAGAAC 44916
QY 41 AapglyArgAlaserTTPglnlnlnalalAAspargleuThrAspleuValalagluArg 60
DB 44917 GACGGGGGGGCTTCTCGGACGAGCGCGGACCGGCTCACCGGACCGGACCGGACCGG 44976

```

QY 61 ThrValLeuAspGlyValValArgLeuLeuAspValGlyCysGlyThrGlyLeuProAla 80
|
|
|
Db 44977 ACCGCTCTCGATGGCGGCTTGACTGCATGTGGGGTGGGTACCGGACMACACGCG 45036
|
|
|
QY 81 LeuArgValAlaArgAspAlaIleGlnIleThrGlyIleThrValSerGlnValGln 100
|
|
|
Db 45037 CTGGCGCTGCGCGGACACCGCATCCAGATCACCGGCATCACCGTCAACCGAGTGCNA 45096
|
|
|
QY 101 ValAlaIleAlaAlaAspCysAlaArgGluArgGlyLeuSerHisArgValAspPheSer 120
|
|
|
Db 45097 GTGGCATGCGCGCATGATGCCACGCGAAACCGGACTTAAGCCACCGGGTGACTTCTCG 45156
|
|
|
QY 121 CysValAspAlaMetSerLeuProTyrProAspAlaIlePheAspAlaAlaTrrAlaMet 140
|
|
|
Db 45157 TCGCGTGAATGATGTCCCTGCGTACCAGCAATGCTTTTGACGCCGCTGGGCCATG 45216
|
|
|
QY 141 GlnSerLeuLeuGluMetSerGluProAspArgAlaIleArgGluIleLeuArgValLeu 160
|
|
|
Db 45217 CAGTGGCTTGGAGATGTCCAGACCGGACCGTCCATCCGGAAATCCTTCGAGTACTC 45276
|
|
|
QY 161 LysProGlyGlyIleLeuGlyValThrGluValValLysArgGluAlaGlyGlyMet 180
|
|
|
Db 45277 AAACCGGATGCGCATCTCGCGCTCACCGAGTCTCAAAACGAGAACGGGCGCGGATG 45336
|
|
|
QY 181 ProValSerGlyAspArgTrrProThrGlyLeuArgIleCysLeuAlaGluGlnLeu 200
|
|
|
Db 45337 CCGGATCGCGGAGACAGGGGCGGCTTCGAGATCCTCGATTGGAGAGCTGTCCGCGAGACC 45396
|
|
|
QY 201 GlnSerLeuArgAlaAlaGlyPheGluIleLeuAspTrrGluAspValSerSerArgThr 220
|
|
|
Db 45397 GAATGGCTCCGCGACGGGGTTCGAGATCCTCGATTGGAGAGCTGTCCGCGAGACC 45456
|
|
|
QY 221 ArgTyrPheMetProGlnPheAlaGluGluLeuAlaAlaHisGlnHisGlyIleAlaAsp 240
|
|
|
Db 45457 CGGTACTTCATGCCACAGTTCGCCAAGAGCTCGCTGGCCACACGCGATCCCGGAC 45516
|
|
|
QY 241 ArgTyrGlyProAlaValAlaGlyTrrAlaAlaAlaValCysAspTrrGluSerTrrAla 260
|
|
|
Db 45517 AGGTACGGGCGCGCTGTCCGCGCTGGCGCGCGGCTGTGGATTATGAGAAATATGCC 45576
|
|
|
QY 261 HisAspMetGlyTrrAlaIleLeuThrAlaArgLysProValGly 275
|
|
|
Db 45577 CACGACATGGGCTATGCAATTCTGAACGGGCGGAAACCGGTGCGC 45621
|
|
|
RESULT 4
AAZ21501
ID AAZ21501 standard; DNA; 80161 BP.
XX
AC AAZ21501;
XX
DT 01-DEC-1999 (first entry)
XX
DE DNA fragment of Saccharopolyspora spinosa containing biosynthetic genes.
XX
KW Spirosyn biosynthetic enzyme; open reading frame; ORF;
KW insecticidal microides; arachnid; nematode; insect; polyketide;
KW polyketide synthase; PKS; extender module; initiator module;
KW acyl transferase domain; AT; acyl carrier protein; ACP;
KW beta-ketosynthase domain; KS; KR; dehydratase domain; DH;
KW enoyl reductase domain; ER; beta-ketoreductase; insecticide; ss.
OS Saccharopolyspora spinosa.
XX
FH Key Location/Qualifiers
FT CDS complement(1135..1971)
FT /tag= a
FT /product= "ORFL16"
FT /note= "Protein involved in transcription control"
FT 2024..2791
FT /tag= b
FT /product= "ORFL15"
FT /note= "Keto acyl reductase"

FT CDS complement(3416..4165)
FT /tag= c
FT /product= "spnS"
FT /note= "Spirosyn biosynthesis protein S"
FT complement(4168..5325)
FT /tag= d
FT /product= "spnR"
FT /note= "Spirosyn biosynthesis protein R"
FT complement(5363..6751)
FT /tag= e
FT /product= "spnQ"
FT /note= "Spirosyn biosynthesis protein Q"
FT 7083..8450
FT /tag= f
FT /product= "spnP"
FT /note= "Spirosyn biosynthesis protein P"
FT 8967..10427
FT /tag= g
FT /product= "SpnO"
FT /note= "Spirosyn biosynthesis protein O"
FT 10436..11434
FT /tag= h
FT /product= "spnM"
FT /note= "Spirosyn biosynthesis protein M"
FT complement(11530..12492)
FT /tag= i
FT /product= "spnM"
FT /note= "Spirosyn biosynthesis protein M"
FT complement(12696..13547)
FT /tag= j
FT /product= "spnL"
FT /note= "Spirosyn biosynthesis protein L"
FT complement(13592..14785)
FT /tag= k
FT /product= "spnK"
FT /note= "Spirosyn biosynthesis protein K"
FT complement(14799..16418)
FT /tag= l
FT /product= "spnJ"
FT /note= "Spirosyn biosynthesis protein J"
FT 16556..17743
FT /tag= m
FT /product= "spnI"
FT /note= "Spirosyn biosynthesis protein I"
FT complement(17749..18501)
FT /tag= n
FT /product= "spnH"
FT /note= "Spirosyn biosynthesis protein H"
FT complement(18541..19713)
FT /tag= o
FT /product= "spnG"
FT /note= "Spirosyn biosynthesis protein G"
FT 20168..20995
FT /tag= p
FT /product= "spnF"
FT /note= "Spirosyn biosynthesis protein F"
FT 21111..28898
FT /tag= q
FT /product= "spnA"
FT /note= "Spirosyn biosynthesis protein A"
FT /transl_except= (Pos:26940..26942, aa:Pro)
FT 21126..22379
FT /tag= r
FT /note= "Beta-ketosynthase domain (KSI): part of the
FT initiator module"
FT 22692..23692
FT /tag= s
FT /product= "spnA"
FT /note= "Acyl transferase domain (ATL): part of the
FT initiator module"
FT 23793..24041
FT /tag= t
FT /product= "Acyl carrier protein domain (ACPI): part of the
FT initiator module"

```
FT misc_feature 24102..25349
FT /tag= u
FT /note= "Beta-ketosynthase domain (KS1) : part of extender
FT module 1"
FT misc_feature 25683..26684
FT /tag= v
FT /note= "Acyl transferase domain (AT1) : part of extender
FT module 1"
FT misc_feature 27582..28121
FT /tag= w
FT /note= "Beta-ketoreductase domain (KR1) : part of extender
FT module 1"
FT misc_feature 28404..28649
FT /tag= x
FT /note= "Acyl carrier protein domain (ACP1) : part of
FT extender module 1"
FT misc_feature 29024..30295
FT /tag= z
FT /note= "Beta-ketosynthase domain (KS2) : part of extender
FT module 2"
FT misc_feature 30629..31621
FT /tag= aa
FT /note= "Acyl transferase domain (AT2) : part of extender
FT module 2"
FT misc_feature 31697..32254
FT /tag= ab
FT /note= "Dehydratase domain (DH2) : part of extender module
FT 2"
FT misc_feature 33035..34072
FT /tag= ac
FT /note= "Enoyl reductase domain (ER2) : part of extender
FT module 2"
FT misc_feature 34082..34621
FT /tag= ad
FT /note= "Beta-ketoreductase domain (KR2) : part of extender
FT module 2"
FT misc_feature 34886..30295
FT /tag= ae
FT /note= "Acyl carrier protein domain (ACP2) : part of
FT extender module 2"
FT CDS 35419..44931
FT /tag= af
FT /product= "spnc"
FT /note= "Spinosyn biosynthesis protein C"
FT 35518..36786
FT /tag= ag
FT /note= "Beta-ketosynthase domain (KS3) : part of extender
FT module 3"
FT misc_feature 37108..38097
FT /tag= ah
FT /note= "Acyl transferase domain (AT3) : part of extender
FT module 3"
FT CDS 38916..35374
FT /tag= ai
FT /product= "spnc"
FT /note= "Spinosyn biosynthesis protein C"
FT 38992..39528
FT /tag= aj
FT /note= "Beta-ketoreductase domain (KR3) : part of extender
FT module 3"
FT misc_feature 39790..40035
FT /tag= aj
FT /note= "Acyl carrier protein domain (ACP3) : part of
FT extender module 3"
FT misc_feature 40102..41373
FT /tag= ak
FT /note= "Beta-ketosynthase domain (KS4) : part of extender
FT module 4"
FT misc_feature 41713..42705
FT /tag= al
FT /note= "Acyl transferase domain (AT4) : part of extender
FT module 4"
FT misc_feature 43615..44157
```

```
FT /tag= am
FT /note= "Beta-ketoreductase domain (KR4) : part of extender
FT module 4"
FT misc_feature 44431..44676
FT /tag= an
FT /note= "Acyl carrier protein domain (ACP4) : part of
FT extender module 4"
FT CDS 44966..59752
FT /tag= ao
FT /product= "spnd"
FT /note= "Spinosyn biosynthesis protein D"
FT 45077..46348
FT misc_feature 46348
FT /tag= ap
FT /note= "Beta-ketosynthase domain (KS5) : part of extender
FT module 5"
FT misc_feature 46691..47674
FT /tag= ap
FT /note= "Acyl transferase domain (AT5) : part of extender
FT module 5"
FT misc_feature 47753..48310
FT /tag= ar
FT /note= "Dehydratase domain (DH5) : part of extender module
FT 5"
FT misc_feature 49226..49771
FT /tag= as
FT /note= "Beta-ketoreductase domain (KR5) : part of extender
FT module 5"
FT misc_feature 50009..50254
```

Alignment Scores:

```
Prod. No.: 2,966-263 Length: 80161
Score: 275.00 Matches: 275
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0
```

US-10-069-353a-8 (1-275) x AAZ21501 (1-80161)

```
QY 1 ValLeuProGlyIValAProThSerGInGInValIglYImetTyRAspleuValThr 20
DB 20168 GTTGGCCAGTGGGCGCACCAACATCGACGAGTGGGCGAGATGATGATGACCTGGTCAAG 20227
QY 21 ProLeuLeuAenSerValAlaGlyIProCySaIaIleHisIGLYTyrTrpGluAen 40
DB 20228 CCGTTCGTAACCTCGATCGCGGCGCGCCCTGCGGCATCCACGAGGACTGAGGAGAAC 20287
QY 41 AspGlyATsGAlaSerTrpGInGInAlaAlaAspATgLeuThrAspLeuValAlaGluArg 60
DB 20288 GACGGGCGGAGCTTCTCTGCGACAGCGCCCGGACCGGCTCACCGACTTGTCCCGAACCG 20347
QY 61 ThrValLeuAspGlyIValIArgLeuLeuAspValIGlyCyGlyThrGlyInProAla 80
DB 20348 ACCGCTCTGATGGCGCGCTTGCATCGCTGATGAGGATGAGGATGAGGATGAGGATGAGG 20407
QY 81 LeuAIGValAlaArgAspAenAlaIleGInIleThrGlyIleThrValSerGInValGIn 100
DB 20408 CTGCGGCTGCGCGCGGACCAACCGATCCAGATCACCGGATCACCGGATCACCGGATCAC 20467
QY 101 ValAlaIleAlaAlaAspCySaIaArgGluATgGlyLeuSerHisArgValAspPheSer 120
DB 20468 GTGGCCATGCGCGCTGATGGCGCGACCGGACCGGACCGGATGAGGATGAGGATGAGGAT 20527
QY 121 CySaValAspAlaMetSerLeuProTyRProAspAenAlaPheAspAlaAlaTrpAlaMet 140
DB 20528 TCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 20587
QY 141 GlnSerLeuLeuGluMetSerGluProAspAenAlaIleArgGluIleLeuArgValLeu 160
DB 20588 CAGTGGCTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 20647
QY 161 LysProGlyIValIleGluGlyValThrGluValValIlyAspGlyAlaGlyIValMet 180
```



```

Db      20648 AACCCGGTGCATCCTCGGCGGCAACCGAGTGTCTCAAAACGAGAACCGGCGCGGATG 20707
Qy      181 ProValSerGlyAspArgTrpProThrGlyLeuAglIecyleuAlaGluGlnleu 200
Db      20708 CCGGGTCCGGGACAGCGTGGCGACCGGCTTCGATTCCTGCTGACCACTTCG 20767
Qy      201 GluSerLeuArgAlaAlaGlyPheGluIleLeuAspTrpGluAspValSerSerArgThr 220
Db      20768 GAATGCTCCGTCGACCGGGGTTCCAGATCTCGATTGGAGAGACTGTCTCGAGGACC 20827
Qy      221 ArgTyrPheMetProGlnPheAlaGluGluLeuAlaAlaGlnIleGlyTLea1aAsp 240
Db      20828 CCGTACTTCATGCCCCAGATTCCCGAAGAGCTCGCTGCCACACAGCAGCGGATCCGGAGC 20887
Qy      241 ArgTyrGlyProAlaValAlaGlyTTPraAlaAlaValaCysAspTyrGluLysTyrAla 260
Db      20888 AGGTACGGGCGCGCTGTCCGCGCGCGCGGCTGTGGATTATGAGAAATATGCC 20947
Qy      261 HisAspMetGlyTyrAlaIleLeuThrAlaArgLysProValGly 275
Db      20948 CACGACATGGCTATCGCATTTCTGACGCGCGGAAACCGGTCGCC 20992

RESULT 5
ABV75558/c
ID      ABV75558 standard; DNA; 36538 BP.
XX
AC      ABV75558;
XX
DT      22-JAN-2003 (first entry)
XX
DE      Saccharopolyspora butenyl-spinosyn biosynthetic gene cluster 2.
XX
KM      Butenyl; biosynthetic enzyme; PKS; polyketide synthetase; macrolide;
metabolite; spinosyn; gene; ds.
XX
OS      Saccharopolyspora sp.
XX
Key     Location/Qualifiers
CDS     complement(114..938)
        /tag= a
        /product= "busF"
        /note= "No start codon given"
        /tag= b
        /product= "busG"
        /tag= c
        /product= "busH"
        complement(3359..4546)
        /tag= d
        /product= "busI"
        /tag= e
        /product= "busJ"
        /tag= f
        /product= "busK"
        /tag= g
        /product= "busL"
        /note= "No start codon given"
        /tag= h
        /product= "busM"
        /note= "No start codon given"
        complement(9668..10666)
        /tag= i
        /product= "busN"
        complement(10675..12135)
        /tag= j
        /product= "busO"
        complement(12864..14177)
        /tag= k

```

```

FT      /product= "busP"
FT      14627..15970
FT      /tag= l
        /product= "busQ"
        /note= "No start codon given"
        /tag= m
        /product= "busR"
        /note= "No start codon given"
        /tag= n
        /product= "busS"
        /tag= o
        /product= "ORF LI"
        complement(19520..19932)
        /tag= p
        /product= "ORF LII"
        /note= "No start codon given"
        complement(20536..21033)
        /tag= q
        /product= "ORF LIII"
        /note= "No start codon given"
        /tag= r
        /product= "ORF LIV"
        complement(22671..23453)
        /tag= s
        /product= "ORF LVI"
        complement(23687..24886)
        /tag= t
        /product= "ORF LVII"
        complement(26177..26923)
        /tag= u
        /product= "ORF LVIII"
        /note= "No start codon given"
        /tag= v
        /product= "ORF LIX"
        /note= "No start codon given"

CDS     WO200279477-A2.
        10-OCT-2002.
        28-MAR-2002; 2002WO-US009968.
        30-MAR-2001; 2001US-0280175P.
        (DOWC ) DOW AGROSCIENCES LLC.
        Hahn DR, Jackson JD, Bullard BS, Gustafson GD, Waldron C;
        Mitchell JC;
        WPI; 2003-058434/05.
        DR P-PDSB; ABP57683, ABP57684, ABP57685, ABP57686, ABP57687, ABP57688,
        DR ABP57689, ABP57690, ABP57691, ABP57692, ABP57693, ABP57694, ABP57695,
        DR ABP57696, ABP57697, ABP57698, ABP57699, ABP57700, ABP57701, ABP57702,
        DR ABP57703, ABP57704.
        New butenyl-spinosyn biosynthetic gene, useful for increasing the
        production of butenyl-spinosyn insecticidal macrolides, or for changing
        the metabolites or products produced by spinosyn-producing
        microorganisms.
        Claim 2; Page 99-119; 218pp; English.
        The invention relates to a novel DNA molecule comprising a DNA sequence
        that encodes a butenyl-spinosyn biosynthetic enzyme, a butenyl-spinosyn
        PKS (polyketide synthetase) domain, or a spinosyn PKS module. The butenyl-
        spinosyn biosynthetic genes are useful for increasing the production of
        butenyl-spinosyn insecticidal macrolides. The genes are also useful for
        changing the metabolites or products produced by spinosyn-producing

```

CC microorganisms. The present sequence represents a DNA molecule encoding
CC butenyl-sphingosyl biosynthetic enzymes

XX
SQ Sequence 36538 BP; 6867 A; 12266 C; 11182 G; 6223 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,51e-25	Length:	36538
Score:	37.00	Matches:	37
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	13.45%	Indels:	0
DB:	10	Gaps:	0

US-10-069-353A-8 (1-275) x ABV75558 (1-36538)

QY 7 ProthSergInGlnValGlyInMetTYrAspLeuValThrProLeuLeuSerVal 26
DB 920 CCAACATCCAGAGATCCGGCAGATGTATGACTGTACCCCGTTGCTGATTCGTC 861

QY 27 AAGAGLYGlyProCysAlaIleHisGlyTYrTyrGluSerAspGlyArg 43
DB 860 GCGGCGCGCCCGCCGATCCACACGCTACTGGAGAACGACGCGCGG 810

RESULT 6
ACL2474/c
ID ACL2474 standard; DNA; 351 BP.

XX
AC ACL2474;

XX
DT 27-OCT-2003 (revised)

XX
DT 17-OCT-2003 (first entry)

XX
DE DNA clone originating in barley containing SNP encoding sequence #12465.

XX
KM Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;

XX
OS Hordelum vulgare; var. (cul.Haruna Nijo).

XX
PN MO2003057877-A1.

XX
PD 17-UTL-2003.

XX
PF 16-DEC-2002; 2002WO-1B005403.

XX
PR 20-DEC-2001; 2001JP-00387059.

XX
PR 20-DEC-2001; 2001JP-00387131.

XX
PR 20-DEC-2001; 2001JP-00403299.

XX
PR 20-DEC-2001; 2001JP-00403300.

XX
PR 27-SEP-2002; 2002JP-00327515.

XX
PA (UYNI-) UNIV JAPAN OKAYAMA.

XX
PI Sato K, Takeda K, Kohara Y;

XX
DR WPI; 2003-587127/55.

XX
PT Single nucleotide polymorphism sites in barley varieties and DNA

XX
PT sequences containing them for analysis and identification of barley

XX
PT varieties and production of barley transformants with desired

XX
PT characteristics.

XX
PS Disclosure; SEQ ID XX; 284pp; Japanese.

CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)

XX
SQ Sequence 351 BP; 95 A; 84 C; 64 G; 108 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.87	Length:	351
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.64%	Indels:	0
DB:	9	Gaps:	0

US-10-069-353A-8 (1-275) x ACL2474 (1-351)

QY 148 GIUPROAspArgAlaIleArgGluIleLeu 157
DB 169 GAACCGGATCGTGCATTCGGAGATTTTA 140

RESULT 7
ACL2489/c
ID ACL2489 standard; DNA; 397 BP.

XX
AC ACL2489;

XX
DT 27-OCT-2003 (revised)

XX
DT 17-OCT-2003 (first entry)

XX
DE DNA clone originating in barley containing SNP encoding sequence #12480.

XX
KM Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;

XX
KM gene; ss.

XX
OS Hordelum vulgare; var. (cul.Haruna Nijo).

XX
PN MO2003057877-A1.

XX
PD 17-UTL-2003.

XX
PF 16-DEC-2002; 2002WO-1B005403.

XX
PR 20-DEC-2001; 2001JP-00387059.

XX
PR 20-DEC-2001; 2001JP-00387131.

XX
PR 20-DEC-2001; 2001JP-00403299.

XX
PR 20-DEC-2001; 2001JP-00403300.

XX
PR 27-SEP-2002; 2002JP-00327515.

XX
PA (UYNI-) UNIV JAPAN OKAYAMA.

XX
PI Sato K, Takeda K, Kohara Y;

XX
DR WPI; 2003-587127/55.

XX
PT Single nucleotide polymorphism sites in barley varieties and DNA

XX
PT sequences containing them for analysis and identification of barley

XX
PT varieties and production of barley transformants with desired

XX
PT characteristics.

XX
PS Disclosure; SEQ ID XX; 284pp; Japanese.

```
XX
SQ Sequence 397 BP; 105 A; 92 C; 70 G; 130 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.1 Length: 397
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.64% Indels: 0
DB: Gaps: 0

US-10-069-353a-8 (1-275) x ACL22489 (1-397)

Qy 148 GIUPROASPARGALAIleArgGluIleLeu 157
Db 238 GAACCGATCGTGGATTGGAGATTTTA 209

RESULT 8
ACL22477/c
ID ACL22477 standard; DNA; 406 BP.
XX
AC ACL22477;
XX
DT 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)
XX
DE DNA clone originating in barley containing SNP encoding sequence #12468.
XX
KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.
XX
OS Hordeum vulgare; var. (cul.Haruna Nijo).
XX
PN WO2003057877-A1.
XX
PD 17-JUL-2003.
XX
PF 16-DEC-2002; 2002WO-IB005403.
XX
PR 20-DEC-2001; 2001JP-00387059.
PR 20-DEC-2001; 2001JP-00387131.
PR 20-DEC-2001; 2001JP-00403299.
PR 20-DEC-2001; 2001JP-00403300.
PR 27-SEP-2002; 2002JP-00327515.
XX
PA (UYN1-) UNIV JAPAN OKAYAMA.
XX
PI Sato K, Takeda K, Kohara Y;
XX
DR WPI; 2003-587127/55.
XX
PT Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.
XX
PS Disclosure; SEQ ID XX; 284pp; Japanese.
XX
CC The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 406 BP; 108 A; 96 C; 74 G; 127 T; 0 U; 1 Other;
```

```
Alignment Scores:
Pred. No.: 2.15 Length: 406
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.64% Indels: 0
DB: Gaps: 0

US-10-069-353a-8 (1-275) x ACL22477 (1-406)

Qy 148 GIUPROASPARGALAIleArgGluIleLeu 157
Db 168 GAACCGATCGTGGATTGGAGATTTTA 139

RESULT 9
ACL22483/c
ID ACL22483 standard; DNA; 439 BP.
XX
AC ACL22483;
XX
DT 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)
XX
DE DNA clone originating in barley containing SNP encoding sequence #12474.
XX
KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.
XX
OS Hordeum vulgare; var. (cul.Haruna Nijo).
XX
PN WO2003057877-A1.
XX
PD 17-JUL-2003.
XX
PF 16-DEC-2002; 2002WO-IB005403.
XX
PR 20-DEC-2001; 2001JP-00387059.
PR 20-DEC-2001; 2001JP-00387131.
PR 20-DEC-2001; 2001JP-00403299.
PR 20-DEC-2001; 2001JP-00403300.
PR 27-SEP-2002; 2002JP-00327515.
XX
PA (UYN1-) UNIV JAPAN OKAYAMA.
XX
PI Sato K, Takeda K, Kohara Y;
XX
DR WPI; 2003-587127/55.
XX
PT Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.
XX
PS Disclosure; SEQ ID XX; 284pp; Japanese.
XX
CC The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 439 BP; 117 A; 106 C; 81 G; 134 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 2.31 Length: 439
Score: 10.00 Matches: 10
```

Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.64% Indels: 0
 DB: 9 Gaps: 0

US-10-069-353A-8 (1-275) x ACL22483 (1-439)

OY 148 GIUProAspArgAlaIleArgGluIleLeu 157
 Db 153 GAACCGATCGTCGATTCGGAGATTTTA 124

RESULT 10

ACL22491/C
 ID ACL22491 standard; DNA; 552 BP.

XX ACL22491;

XX 27-OCT-2003 (revised)

DT 17-OCT-2003 (first entry)

XX DNA clone originating in barley containing SNP encoding sequence #12482.

DE Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;

XX gene; ss.

KW Hordeum vulgare; var. (cul.Haruna Ni'jo).

XX WO2003057877-A1.

XX 17-JUL-2003.

PF 16-DEC-2002; 2002WO-1B005403.

XX 20-DEC-2001; 2001JP-00387059.

PR 20-DEC-2001; 2001JP-00387131.

PR 20-DEC-2001; 2001JP-00403299.

PR 20-DEC-2001; 2001JP-00403300.

PR 27-SEP-2002; 2002JP-00327515.

XX (UYNI-) UNIV JAPAN OKAYAMA.

XX Sato K, Takeda K, Kohara Y;

XX MPI; 2003-587127/55.

XX Single nucleotide polymorphism sites in barley varieties and DNA

PT sequences containing them for analysis and identification of barley

PT varieties and production of barley transformants with desired

PT characteristics.

XX Disclosure; SEQ ID XX; 284pp; Japanese.

XX The present invention relates to oligonucleotide clones originating in

CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms

CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley

CC varieties, identification of particular varieties and genotype-phenotype

CC analysis, isolation of specific genes and creation of new varieties by

CC transformation of barley varieties with them and production of new barley

CC varieties with desired properties. The present sequence represents an

CC oligonucleotide clone DNA sequence featured in the specification. The

CC sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at fcp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to

CC standardise OS field)

XX Sequence 552 BP; 147 A; 137 C; 102 G; 165 T; 0 U; 1 Other;

SO Alignment Scores:

Pred. No.: 2.87 Length: 552

Score: 10.00 Matches: 10

Percent Similarity: 100.00% Conservativity: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 3.64% Indels: 0

DB: 9 Gaps: 0

US-10-069-353A-8 (1-275) x ACL22491 (1-552)

OY 148 GIUProAspArgAlaIleArgGluIleLeu 157
 Db 201 GAACCGATCGTCGATTCGGAGATTTTA 172

RESULT 11

ACL22487/C
 ID ACL22487 standard; DNA; 572 BP.

XX ACL22487;

XX 27-OCT-2003 (revised)

DT 17-OCT-2003 (first entry)

XX DNA clone originating in barley containing SNP encoding sequence #12478.

DE Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;

XX gene; ss.

KW Hordeum vulgare; var. (cul.Haruna Ni'jo).

XX WO2003057877-A1.

XX 17-JUL-2003.

PF 16-DEC-2002; 2002WO-1B005403.

XX 20-DEC-2001; 2001JP-00387059.

PR 20-DEC-2001; 2001JP-00387131.

PR 20-DEC-2001; 2001JP-00403299.

PR 20-DEC-2001; 2001JP-00403300.

PR 27-SEP-2002; 2002JP-00327515.

XX (UYNI-) UNIV JAPAN OKAYAMA.

XX Sato K, Takeda K, Kohara Y;

XX MPI; 2003-587127/55.

XX Single nucleotide polymorphism sites in barley varieties and DNA

PT sequences containing them for analysis and identification of barley

PT varieties and production of barley transformants with desired

PT characteristics.

XX Disclosure; SEQ ID XX; 284pp; Japanese.

XX The present invention relates to oligonucleotide clones originating in

CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms

CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley

CC varieties, identification of particular varieties and genotype-phenotype

CC analysis, isolation of specific genes and creation of new varieties by

CC transformation of barley varieties with them and production of new barley

CC varieties with desired properties. The present sequence represents an

CC oligonucleotide clone DNA sequence featured in the specification. The

CC sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at fcp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to

CC standardise OS field)

XX Sequence 572 BP; 151 A; 145 C; 110 G; 166 T; 0 U; 0 Other;

SO Alignment Scores:

Pred. No.: 2.97 Length: 572

Score: 10.00 Matches: 10

Percent Similarity: 100.00% Conservativity: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 3.64% Indels: 0

DB: 9 Gaps: 0

US-10-069-353A-8 (1-275) x ACL22487 (1-572)

```
OY 148 GIUProaSPARGAla1leArgGluIleLeu 157
Db 201 GAACCGATCGTCGATTCGGAGATTTTA 172
RESULT 12
AC122468/C
ID AC122468 standard; DNA; 607 BP.
XX
XX AC122468;
XX
XX 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)
XX
XX DNA clone originating in barley containing SNP encoding sequence #12459.
DE
XX
XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.
XX
XX Hordeum vulgare; var. (cul.Haruna Nijo).
OS
XX
XX WO2003057877-A1.
XX
XX 17-JUL-2003.
XX
XX 16-DEC-2002; 2002WO-1B005403.
XX
XX PF 20-DEC-2001; 2001JP-00387059.
XX
XX PR 20-DEC-2001; 2001JP-00387131.
XX
XX PR 20-DEC-2001; 2001JP-00403299.
XX
XX PR 20-DEC-2001; 2001JP-00403300.
XX
XX PR 27-SEP-2002; 2002JP-00327515.
XX
XX
XX (UNNI-) UNIV JAPAN OKAYAMA.
XX
XX PI Sato K, Takeda K, Kohara Y;
XX
XX DR WPI; 2003-587127/55.
XX
XX PT Single nucleotide polymorphism sites in barley varieties and DNA
XX PT sequences containing them for analysis and identification of barley
XX PT varieties and production of barley transformants with desired
XX PT characteristics.
XX
XX PS Disclosure; SEQ ID XX; 284bp; Japanese.
XX
XX CC The present invention relates to oligonucleotide clones originating in
XX CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
XX CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
XX CC varieties, identification of particular varieties and genotype-phenotype
XX CC analysis, isolation of specific genes and creation of new varieties by
XX CC transformation of barley varieties with them and production of new barley
XX CC varieties with desired properties. The present sequence represents an
XX CC oligonucleotide clone DNA sequence featured in the specification. The
XX CC sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
XX CC standardise OS field)
XX
XX SQ Sequence 607 BP; 159 A; 148 C; 126 G; 174 T; 0 U; 0 Other;
XX
XX
Alignment Scores:
Pred. No.: 3.14 Length: 607
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.64% Indels: 0
DB: 9 Gaps: 0
US-10-069-353A-8 (1-275) x AC122468 (1-607)
OY 148 GIUProaSPARGAla1leArgGluIleLeu 157
|||||
|||||
```

```
DB 171 GAACCGATCGTCGATTCGGAGATTTTA 142
RESULT 13
AC122454/C
ID AC122454 standard; DNA; 617 BP.
XX
XX AC122454;
XX
XX 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)
XX
XX DNA clone originating in barley containing SNP encoding sequence #12445.
DE
XX
XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.
XX
XX Hordeum vulgare; var. (cul.Haruna Nijo).
OS
XX
XX WO2003057877-A1.
XX
XX 17-JUL-2003.
XX
XX 16-DEC-2002; 2002WO-1B005403.
XX
XX PF 20-DEC-2001; 2001JP-00387059.
XX
XX PR 20-DEC-2001; 2001JP-00387131.
XX
XX PR 20-DEC-2001; 2001JP-00403299.
XX
XX PR 20-DEC-2001; 2001JP-00403300.
XX
XX PR 27-SEP-2002; 2002JP-00327515.
XX
XX
XX (UNNI-) UNIV JAPAN OKAYAMA.
XX
XX PI Sato K, Takeda K, Kohara Y;
XX
XX DR WPI; 2003-587127/55.
XX
XX PT Single nucleotide polymorphism sites in barley varieties and DNA
XX PT sequences containing them for analysis and identification of barley
XX PT varieties and production of barley transformants with desired
XX PT characteristics.
XX
XX PS Disclosure; SEQ ID XX; 284bp; Japanese.
XX
XX CC The present invention relates to oligonucleotide clones originating in
XX CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
XX CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
XX CC varieties, identification of particular varieties and genotype-phenotype
XX CC analysis, isolation of specific genes and creation of new varieties by
XX CC transformation of barley varieties with them and production of new barley
XX CC varieties with desired properties. The present sequence represents an
XX CC oligonucleotide clone DNA sequence featured in the specification. The
XX CC sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
XX CC standardise OS field)
XX
XX SQ Sequence 617 BP; 160 A; 154 C; 125 G; 178 T; 0 U; 0 Other;
XX
XX
Alignment Scores:
Pred. No.: 3.19 Length: 617
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.64% Indels: 0
DB: 9 Gaps: 0
US-10-069-353A-8 (1-275) x AC122454 (1-617)
OY 148 GIUProaSPARGAla1leArgGluIleLeu 157
Db 190 GAACCGATCGTCGATTCGGAGATTTTA 161
RESULT 14
|||||
|||||
```

```

ACL22486/c
ID  ACL22486 standard; DNA; 632 BP.
XX
AC  ACL22486;
XX
DT  27-OCT-2003 (revised)
XX
DT  17-OCT-2003 (first entry)
XX
DE  DNA clone originating in barley containing SNP encoding sequence #12477.
XX
KM  Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
XX  gene; ss.
XX
OS  Hordeum vulgare; var. (cul.Haruna Nijo).
XX
PN  WO2003057877-A1.
XX
PD  17-JUL-2003.
XX
PF  16-DEC-2002; 2002WO-IB005403.
XX
PR  20-DEC-2001; 2001JP-00387059.
XX  PR  20-DEC-2001; 2001JP-00387131.
XX  PR  20-DEC-2001; 2001JP-00403299.
XX  PR  20-DEC-2001; 2001JP-00403300.
XX  PR  27-SEP-2002; 2002JP-00327515.
XX
PA  (UYNI-) UNIV JAPAN OKAYAMA.
XX
PI  Sato K, Takeda K, Kohara Y;
XX  WPI; 2003-587127/55.
XX
PT  Single nucleotide polymorphism sites in barley varieties and DNA
XX  PT  sequences containing them for analysis and identification of barley
XX  PT  varieties and production of barley transformants with desired
XX  PT  characteristics.
XX
PS  Disclosure; SEQ ID XX; 284bp; Japanese.
XX
CC  The present invention relates to oligonucleotide clones originating in
XX  CC  barley (Hordeum vulgare) which contain single nucleotide polymorphisms
XX  CC  (SNP). The oligonucleotides may be used for analysis of SNPs among barley
XX  CC  varieties, identification of particular varieties and genotype-phenotype
XX  CC  analysis, isolation of specific genes and creation of new varieties by
XX  CC  transformation of barley varieties with them and production of new barley
XX  CC  varieties with desired properties. The present sequence represents an
XX  CC  oligonucleotide clone DNA sequence featured in the specification. The
XX  CC  sequence data for this patent did not form part of the printed
XX  CC  specification, but was obtained in electronic format directly from WIPO
XX  CC  at fcp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
XX  CC  standardise OS field)
XX
SQ  Sequence 632 BP; 164 A; 156 C; 130 G; 182 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 3.26 Length: 632
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.64% Indels: 0
DB: 9 Gaps: 0
US-10-069-353a-8 (1-275) x ACL22486 (1-632)
OY  148 GluProAspArgAlaIleArgGluIleLeu 157
Db  193 GAACCGATCGTCGATTCGGGAGATTTTA 164
RESULT 15
ACL22486/c
ID  ACL22486 standard; DNA; 633 BP.
XX

```

```

AC  ACL22484;
XX
DT  27-OCT-2003 (revised)
XX
DT  17-OCT-2003 (first entry)
XX
DE  DNA clone originating in barley containing SNP encoding sequence #12475.
XX
KM  Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
XX  gene; ss.
XX
OS  Hordeum vulgare; var. (cul.Haruna Nijo).
XX
PN  WO2003057877-A1.
XX
PD  17-JUL-2003.
XX
PF  16-DEC-2002; 2002WO-IB005403.
XX
PR  20-DEC-2001; 2001JP-00387059.
XX  PR  20-DEC-2001; 2001JP-00387131.
XX  PR  20-DEC-2001; 2001JP-00403299.
XX  PR  20-DEC-2001; 2001JP-00403300.
XX  PR  27-SEP-2002; 2002JP-00327515.
XX
PA  (UYNI-) UNIV JAPAN OKAYAMA.
XX
PI  Sato K, Takeda K, Kohara Y;
XX  WPI; 2003-587127/55.
XX
PT  Single nucleotide polymorphism sites in barley varieties and DNA
XX  PT  sequences containing them for analysis and identification of barley
XX  PT  varieties and production of barley transformants with desired
XX  PT  characteristics.
XX
PS  Disclosure; SEQ ID XX; 284bp; Japanese.
XX
CC  The present invention relates to oligonucleotide clones originating in
XX  CC  barley (Hordeum vulgare) which contain single nucleotide polymorphisms
XX  CC  (SNP). The oligonucleotides may be used for analysis of SNPs among barley
XX  CC  varieties, identification of particular varieties and genotype-phenotype
XX  CC  analysis, isolation of specific genes and creation of new varieties by
XX  CC  transformation of barley varieties with them and production of new barley
XX  CC  varieties with desired properties. The present sequence represents an
XX  CC  oligonucleotide clone DNA sequence featured in the specification. The
XX  CC  sequence data for this patent did not form part of the printed
XX  CC  specification, but was obtained in electronic format directly from WIPO
XX  CC  at fcp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
XX  CC  standardise OS field)
XX
SQ  Sequence 633 BP; 165 A; 156 C; 130 G; 182 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 3.27 Length: 633
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.64% Indels: 0
DB: 9 Gaps: 0
US-10-069-353a-8 (1-275) x ACL22484 (1-633)
OY  148 GluProAspArgAlaIleArgGluIleLeu 157
Db  194 GAACCGATCGTCGATTCGGGAGATTTTA 165
Search completed: August 29, 2005, 22:45:39
Job time : 673 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 29, 2005, 22:18:55 ; Search time 3540 Seconds

(without alignments)
2956.970 Million cell updates/sec

Title: US-10-069-353A-8

Perfect score: 275

Sequence: 1 VLFGAPATSGQVQGMVDLVT.....YEKXADMGYAILTAKPVG 275

Scoring table:

OLIGO	
Xgapop 60.0 , Xgapext 60.0	
Ygapop 60.0 , Ygapext 60.0	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 34239544 seqs, 19032134700 residues

Word size: 1

Total number of hits satisfying chosen parameters: 68477535

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O/cgnt 1/USPRO.spool/US10069353/runat 26082005 172152 15063/app_query.fasta_1.455
-DB=EST -QMT=fastcap -SUPFIX=ret -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=200000000
-USRR=US10069353 @CGN 1 1 3437 @runat 26082005 172152 15063 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOUTPRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_eest1:*
2: gb_eest2:*
3: gb_hrc:*
4: gb_eest3:*
5: gb_eest4:*
6: gb_eest5:*
7: gb_eest6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	3.6	288	7	CR802983 NF24d11f4
2	10	3.6	324	6	CD662504 UCRV18.0
3	10	3.6	351	1	AV934664 AV934664
4	10	3.6	397	4	BU552337 BU552337
5	10	3.6	406	4	BU548396 BU548396
6	10	3.6	433	4	BG909505 BG909505
7	10	3.6	439	1	AV936770 Talr1104f
8	10	3.6	451	6	CD934354 CD934354
9	10	3.6	504	6	CD881537 CD881537

C 10	10	3.6	511	4	BT50759	BT50759	Taol.09g0
C 11	10	3.6	540	5	BQ620788	BQ620788	Talr1111c
C 12	10	3.6	552	1	BU552981	BU552981	BU552981
C 13	10	3.6	564	6	CD876873	CD876873	AZ03.111C
C 14	10	3.6	567	1	BU459101	BU459101	BU459101
C 15	10	3.6	572	1	AV945873	AV945873	AV945873
C 16	10	3.6	574	4	BU451604	BU451604	BU451604
C 17	10	3.6	576	6	CD862060	CD862060	AZ01.102D
C 18	10	3.6	594	6	CD896883	CD896883	G174.104B
C 19	10	3.6	597	4	BG904766	BG904766	Talr1134G
C 20	10	3.6	607	4	BU550817	BU550817	BU550817
C 21	10	3.6	617	4	BU546846	BU546846	BU546846
C 22	10	3.6	632	1	AV945413	AV945413	AV945413
C 23	10	3.6	633	1	AV942357	AV942357	AV942357
C 24	10	3.6	634	4	BU548698	BU548698	BU548698
C 25	10	3.6	637	6	CA597128	CA597128	WPA1C.pk0
C 26	10	3.6	642	4	BU548260	BU548260	BU548260
C 27	10	3.6	651	4	BU547951	BU547951	BU547951
C 28	10	3.6	653	6	CD934353	CD934353	GR45.123P
C 29	10	3.6	674	4	BU552252	BU552252	BU552252
C 30	10	3.6	685	4	BU548223	BU548223	BU548223
C 31	10	3.6	700	1	AV936690	AV936690	AV936690
C 32	10	3.6	703	4	BU547277	BU547277	BU547277
C 33	10	3.6	703	4	BU548319	BU548319	BU548319
C 34	10	3.6	705	4	BU552839	BU552839	BU552839
C 35	10	3.6	708	1	AV942972	AV942972	AV942972
C 36	10	3.6	719	4	BU546326	BU546326	BU546326
C 37	10	3.6	721	4	BU548517	BU548517	BU548517
C 38	10	3.6	726	4	BU548683	BU548683	BU548683
C 39	10	3.6	734	9	CE354658	CE354658	CE354658
C 40	10	3.6	821	2	BF065038	BF065038	HY.CBD002
C 41	10	3.6	951	2	BF255786	BF255786	BF255786
C 42	9	3.3	139	1	AT061356	AT061356	AC08A03.6
C 43	9	3.3	236	7	CF681341	CF681341	CF681341
C 44	9	3.3	375	1	AU312420	AU312420	AU312420
C 45	9	3.3	397	6	CA735338	CA735338	Wp11b.pk0

ALIGNMENTS

RESULT 1
CR802983 288 bp mRNA linear EST 26-FEB-2004
NF24d11f4.r1 Tail fescue PI297901 44 deg C Heat Stress SSH cDNA
Schedonorus arundinaceus cDNA clone NF24d11f4 5', mRNA sequence.

ACCESSION CR802983
VERSION GI:43404841
KEYWORDS
SOURCE
ORGANISM Schedonorus arundinaceus (Festuca arundinacea)

REFERENCE
Zhang, Y., Zwonitzer, J. C., Chekhovskiy, K., May, G. D., and Mian, M. A. R.
(eds.), MOLECULAR BREEDING OF FORAGE AND TURF, Kluwer Acad. Pub. 0,
1 (bases 1 to 288)

A functional genomics approach for identification of heat tolerance
genes in tall fescue
(in) Hopkins, A., Wang, Z. Y., Mian, R., Sledge, M., and Barker, R. E.
(eds.), MOLECULAR BREEDING OF FORAGE AND TURF, Kluwer Acad. Pub. 0,
87-96 (2003)

COMMENT
Contact: Bruce A. Roe, University of Oklahoma, broeou@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762

FEATURES
source
Contact: Dr. Rouf Mian (rmiannoble.org) regarding clone availability
Seq primer: M13 reverse primer
High quality sequence stop: 281.
Location/Qualifiers
1. .288
/organism="Schedonorus arundinaceus"


```

/mo1_type="mRNA"
/cultivar="PI297901"
/db_xref="taxon:4606"
/clone="NF24d11f44"
/issue_type="shoot"
/lab_host="E.coli"
/clone_1ib="Tail Rescue PI297901 44 deg C Heat Stress SSH
cDNA"
/note="Vector: PCR 2.1-TOPO; Site 1: EcoR I; Site 2: EcoR
I; BD/Clontech PCR-select cDNA subtraction library"

ORIGIN

Alignment Scores:
Pred. No.:      14.1      Length:      288
Score:          10.00     Matches:      10
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:      3.64%      Indels:      0
DB:               7          Gaps:          0

US-10-069-353A-8 (1-275) x CK802983 (1-288)

QY      148 GluProAspArgAlaIleArgGluIleu 157
      |||||
      270 GAACGAGACCGCGCTATTGCGAGATTTG 241

RESULT 2
LOCUS   CD662504               324 bp      mRNA      linear      EST 23-JUN-2003
DEFINITION   UCRHV18_02af09 b1 Drought-stressed Dicktoo barley epidermis cDNA
              library Hordeum vulgare subsp. vulgare cDNA clone UCRHV18_02af09,
              mRNA sequence.
ACCESSION   CD662504
VERSION     CD662504.1 GI:32152807
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare
ORGANISM    Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Hordeum.
REFERENCE   1 (bases 1 to 324)
            Wahid,A., Close,T.J., Fenton,R.D., Manamaker,S., Collura,K.,
            Feuerbacher,O., Kim,H.R., Kudrna,D., Wing,R. and Yu,Y.
            Drought-stressed barley leaf epidermis cDNA sequences
            Unpublished (2003)
            Contact: Timothy J. Close
            Department of Botany & Plant Sciences
            University of California
            Riverside, CA 92521-0124, USA
            Tel: 909-787-3318
            Fax: 909-787-4437
            Email: timothy.close@ucr.edu
            Seq primer: T7
            Location/Qualifiers
                1..324
                /organism="Hordeum vulgare subsp. vulgare"
                /mol_type="mRNA"
                /cultivar="Dicktoo"
                /sub_species="vulgare"
                /db_xref="taxon:112509"
                /clone="UCRHV18_02af09"
                /issue_type="lower leaf epidermis"
                /dev_stage="1-2 week seedlings"
                /lab_host="E. coli TJC121"
                /clone_1ib="Drought-stressed Dicktoo barley epidermis cDNA
                library"
                /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                site 1: EcoRI; Site 2: XhoI; Seeds of barley (Hordeum
                vulgare L. cv. Dicktoo) were germinated in dishes
                containing UCR-mix soil. Seedlings were kept in a growth
                chamber at 20C (day/night) and allowed to grow at 60-70%
                soil moisture content. After 3-4 days, the water was
                withheld in order to apply drought until the soil moisture

```

content was reduced to 10-12%, which took another 3-4 days. At the time of extraction of epidermis, leaf water and osmotic potentials had dropped to 15.9%, 20.2% and 24.7% of controls, respectively, as determined using a vapor pressure osmometer (Model 5100C, Wescor, Inc., Logan, UT). Epidermal strips were quickly peeled off of seedlings and immediately submerged in liquid nitrogen. About 15-20 g of epidermal tissue was collected and used to extract total RNA. Total RNA was extracted using Concert Plant RNA Reagent (Invitrogen 12322-012). Poly(A) RNA was purified using PolyAtrack mRNA Isolation System IV (Promega). A cDNA library was made using a Uni-ZAP cDNA synthesis kit (Stratagene). A total of 1 million primary lambda cDNA clones were mass-excised in vivo to give a population of phagescript SK(-) phagemids. All steps to this point were performed in the T7 Close lab at the University of California, Riverside by A. Wahid with some assistance from R.D. Fenton. Phagemids were plated on the TJC121 host strain, plasmid DNA purified, cDNA clones archived, and DNA sequences determined using the T7 primer (mainly 3' end reads) using an ABI3730 at the Arizona Genomics Institute, University of Arizona (Collura, Feuerbacher, Kim, Kudrna, Wing, Yu). Chromatogram files were transmitted to UC Riverside (by Yu), then processed at UC Riverside (by Manamaker) using the Harvest pipeline (http://harvest.ucr.edu) to remove vector and cloning oligo sequences and various contaminants, and to trim to a high quality region. Sequences that retained a phred 17 region of at least 100 bases were deposited to Genbank."

```

ORIGIN

Alignment Scores:
Pred. No.:      15.8      Length:      324
Score:          10.00     Matches:      10
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:      3.64%      Indels:      0
DB:               6          Gaps:          0

US-10-069-353A-8 (1-275) x CD662504 (1-324)

QY      148 GluProAspArgAlaIleArgGluIleu 157
      |||||
      167 GAACGAGATCGTCGATTCGGAGATTTTA 138

RESULT 3
LOCUS   AV934664               351 bp      mRNA      linear      EST 18-JAN-2002
DEFINITION   AV934664 K. Sato unpublished cDNA library, cv. Hartun NiJo adult,
              heading stage top three leaves Hordeum vulgare subsp. vulgare cDNA
              clone baal1d01 3', mRNA sequence.
ACCESSION   AV934664
VERSION     AV934664.1 GI:18230461
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare
ORGANISM    Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Hordeum.
REFERENCE   1 (bases 1 to 351)
            Sato,K., Saitoh,D. and Takeda,K.
            Barley EST sequencing project in NIG and Okayama Univ
            Unpublished (2002)
            Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshin@genes.nig.ac.jp.
            Location/Qualifiers
                1..351

```

/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Haruna NiJo"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="baa11d01"
/tissue_type="top three leaves"
/dev_stage="adult, heading stage"
/clone_1ib="K. Sato unpublished cDNA library, cv. Haruna NiJo adult, heading stage top three leaves"

ORIGIN

Alignment Scores:

Pred. No.: 17 Length: 351
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.64% Indels: 0
DB: 1 Gaps: 0

US-10-069-353a-8 (1-275) x AV934664 (1-351)

QY 148 GluProAaPARGAlaIleArgGluIleLeu 157
Db 169 GAACCGAGATCGATTCGGAGATTTTA 140

RESULT 4 397 bp mRNA linear EST 14-NOV-2002
LOCUS BJ552337/c
DEFINITION BJ552337 K. Sato unpublished cDNA library, strain H602 adult,
heading stage top three leaves Hordeum vulgare subsp. spontaneum
cDNA clone bah52m15 3', mRNA sequence.

ACCESSION BJ552337 GI:24970788
VERSION BJ552337
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. spontaneum
ORGANISM Hordeum vulgare subsp. spontaneum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.

REFERENCE 1 (bases 1 to 397)
Sato, K., Saisho, D. and Takeda, K.

AUTHORS Barley EST sequencing project in NIG and Okayama Univ
TITLE Unpublished (2002)
JOURNAL Contact: Tadao Shin-i
COMMENT Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.

FEATURES
source 1..397
/organism="Hordeum vulgare subsp. spontaneum"
/mol_type="mRNA"
/strain="H602"
/sub_species="spontaneum"
/db_xref="taxon:77009"
/clone="bah52m15"
/tissue_type="top three leaves"
/dev_stage="adult, heading stage"
/clone_1ib="K. Sato unpublished cDNA library, strain H602 adult, heading stage top three leaves"

ORIGIN

Alignment Scores:

Pred. No.: 19 Length: 397
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.64% Indels: 0
DB: 4 Gaps: 0

US-10-069-353a-8 (1-275) x BJ552337 (1-397)

QY 148 GluProAaPARGAlaIleArgGluIleLeu 157
Db 238 GAACCGAGATCGATTCGGAGATTTTA 209

RESULT 5 406 bp mRNA linear EST 14-NOV-2002
LOCUS BJ548396/c
DEFINITION BJ548396 K. Sato unpublished cDNA library, cv. Haruna NiJo adult,
heading stage top three leaves Hordeum vulgare subsp. vulgare cDNA
clone baa124d16 3', mRNA sequence.

ACCESSION BJ548396 GI:24966834
VERSION BJ548396
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.

REFERENCE 1 (bases 1 to 406)
Sato, K., Saisho, D. and Takeda, K.

AUTHORS Barley EST sequencing project in NIG and Okayama Univ
TITLE Unpublished (2002)
JOURNAL Contact: Tadao Shin-i
COMMENT Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.

FEATURES
source 1..406
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Haruna NiJo"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="baa124d16"
/tissue_type="top three leaves"
/dev_stage="adult, heading stage"
/clone_1ib="K. Sato unpublished cDNA library, cv. Haruna NiJo adult, heading stage top three leaves"

ORIGIN

Alignment Scores:

Pred. No.: 19.4 Length: 406
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.64% Indels: 0
DB: 4 Gaps: 0

US-10-069-353a-8 (1-275) x BJ548396 (1-406)

QY 148 GluProAaPARGAlaIleArgGluIleLeu 157
Db 168 GAACCGAGATCGATTCGGAGATTTTA 139

RESULT 6 433 bp mRNA linear EST 05-JUN-2001
LOCUS BG909505/c
DEFINITION BG909505 Talr1104F08B Talr1 Triticum aestivum cDNA clone Talr1104F08 5',
mRNA sequence.

ACCESSION BG909505 GI:14317181
VERSION BG909505
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.

REFERENCE 1 (bases 1 to 433)

AUTHORS Cloutier, S., Dong, G. and Walsh, A.
 TITLE Wheat functional genomics - Thatcher Lr1 cDNA library
 JOURNAL Unpublished (2001)
 COMMENT Contact: Dr. Sylvie Cloutier
 Cereal Research Centre, Agriculture and Agri-food Canada
 195 Dufferin Rd, Winnipeg, MB, Canada R3T 2M9
 Tel: (204) 983-2340
 Fax: (204) 983-4604
 Email: scloutier@gr.gc.ca

was cloned directionally, not all sequences generated with reverse
 primer were from the 5' end (same with forward primer and 3' end).
 Average insert size is >2.2 kb
 Plate: 104 row: F column: 08
 Seq primer: M13 Reverse.

FEATURES
 source Location/Qualifiers

1..433
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="Thatcher Lr1"
 /db_xref="taxon:4565"
 /clone="TALr1104F08"
 /tissue_type="leaf tissue"
 /dev_stage="14 Days Old"
 /lab_host="E. coli XL0LR"
 /clone_1lb="TALr1"
 /note="Vector: Lambda ZapII; mass excised in plasmid
 vector PBK-CMV (Stratagene); Site 1: EcoRI; Site 2: XhoI;
 mRNA obtained from wheat NIL Thatcher Lr1 24 hours after
 inoculation with leaf rust pathogen Puccinia triticina
 race BBB carrying the avirulence gene Avr1."

ORIGIN

Alignment Scores:

Pred. No.:	20.6	Length:	433
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.64%	Indels:	0
DB:	4	Gaps:	0

US-10-069-353a-8 (1-275) x BG909505 (1-433)

Qy 148 GIUProASPARGALIAIEARGLUILEU 157
 |||||
 Db 237 GAACCGAGTCGCGATTGCGGAGATTG 208

RESULT 7

LOCUS AV936770 439 bp mRNA linear EST 18-JAN-2002
 AV936770 K. Sato unpublished cDNA library, cv. Haruna Nijo adult,
 heading stage top three leaves Hordeum vulgare subsp. vulgare cDNA
 clone baal509 3', mRNA sequence.

ACCESSION AV936770 GI:18232567
 VERSION AV936770.1
 KEYWORDS EST.
 SOURCE Hordeum vulgare subsp. vulgare
 ORGANISM Hordeum vulgare subsp. vulgare

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Hordeum.
 1 (bases 1 to 439)
 Sato, K., Saitoh, D. and Takeda, K.
 TITLE Barley EST sequencing project in NIG and Okayama Univ
 JOURNAL Unpublished (2002)
 COMMENT Contact: Tadao Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@gens.nig.ac.jp.

FEATURES
 source Location/Qualifiers
 1..439

/organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="Haruna Nijo"
 /sub_species="vulgare"
 /db_xref="taxon:112509"
 /clone="baal509"
 /tissue_type="top three leaves"
 /dev_stage="adult, heading stage"
 /clone_1lb="K. Sato unpublished cDNA library, cv. Haruna
 Nijo adult, heading stage top three leaves"

ORIGIN

Alignment Scores:

Pred. No.:	20.8	Length:	439
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.64%	Indels:	0
DB:	1	Gaps:	0

US-10-069-353a-8 (1-275) x AV936770 (1-439)

Qy 148 GIUProASPARGALIAIEARGLUILEU 157
 |||||
 Db 153 GAACCGAGTCGCGATTGCGGAGATTG 124

RESULT 8

LOCUS CD934354 451 bp mRNA linear EST 15-JUL-2003
 CD934354 GR45.123P11R010830 GR45 Triticum aestivum cDNA clone GR45123P11,
 mRNA sequence.

ACCESSION CD934354
 VERSION CD934354.1 GI:32782118
 KEYWORDS EST.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.
 1 (bases 1 to 451)
 Genoplante.
 Genoplante, a major partnership french program in plant genomics
 JOURNAL Unpublished (2003)
 COMMENT Contact: Genoplante
 Genoplante
 93, rue Henri Rochefort 91025 EVRY CEDEX France
 Tel: 33 1 69 47 54 00
 Fax: 33 1 69 47 54 10
 This sequence has been generated in the framework of the french
 plant genomics programme 'Genoplante' (<http://www.genoplante.com>
 and <http://genoplante-info.infobiogen.fr>).

FEATURES
 source Location/Qualifiers
 1..451
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="recital"
 /db_xref="taxon:4565"
 /clone="GR45123P11"
 /tissue_type="grain (45 degrees per day after
 pollination)"
 /clone_1lb="GR45"

ORIGIN

Alignment Scores:

Pred. No.:	21.4	Length:	451
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.64%	Indels:	0
DB:	6	Gaps:	0

US-10-069-353a-8 (1-275) x CD934354 (1-451)

QY 154 ArgGlu11leuArgVal1leuLysProGly 163
 |||||
 Db 388 CGAGAGATCCTTAGAGTCTTAAACACGAGA 359

RESULT 9
 LOCUS CD881537 504 bp mRNA linear EST 14-JUL-2003
 DEFINITION F1.103118F010329 F1 Triticum aestivum cDNA clone F1103118, mRNA
 sequence.
 ACCESSION CD881537
 VERSION CD881537.1 GI:32641050
 KEYWORDS EST
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.
 1 (bases 1 to 504)
 REFERENCE
 AUTHORS Genoplante.
 TITLE Unpublished (2003)
 JOURNAL Contact: Genoplante
 COMMENT Genoplante
 93, rue Henri Rochefort 91025 EVRY CEDEX France
 Tel: 33 1 69 47 54 00
 Fax: 33 1 69 47 54 10
 This sequence has been generated in the framework of the french
 plant genomics programme 'Genoplante' (<http://www.genoplante.com>
 and <http://genoplante-info.infobiogen.fr>).

FEATURES
 source
 1..504
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="recital"
 /db_xref="taxon:4565"
 /clone="F1103118"
 /tissue_type="leaf one"
 /clone_lib="F1"

ORIGIN
 Alignment Scores:
 Pred. No.: 23 7 Length: 504
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.64% Indels: 0
 DB: 6 Gaps: 0

US-10-069-353a-8 (1-275) x CD881537 (1-504)

QY 148 GluProAspArgAla1leuArgGlu1leu 157
 |||||
 Db 267 GAGCCAGACCGTGTATTCCGGAGATTGG 296

RESULT 10
 LOCUS B1750759/c 511 bp mRNA linear EST 25-SEP-2001
 DEFINITION Ta01_09G01_R
 Ta01_AAFc_ECOrc_Fusarium graminearum inoculated wheat heads
 Triticum aestivum cDNA clone Ta01_09G01, mRNA sequence.
 ACCESSION B1750759
 VERSION B1750759.1 GI:15772561
 KEYWORDS EST
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.
 1 (bases 1 to 511)
 REFERENCE
 AUTHORS Ouellet, T., Dan, H., Koul, A., Chapados, V., Couroux, P., De Moors, A.,
 Harrie, L. J., Hattori, J. I., Robert, L. S., Singh, J. A., Sprout, D. and
 Tinker, N. A.
 TITLE Expressed Sequence Tags from Wheat Heads 24 Hours after Spray

JOURNAL
 COMMENT Unpublished (2001)
 Contact: Ouellet, Therese
 Eastern Cereal and Oilseed Research Centre
 Agriculture and Agri-food Canada
 Neaby Bldg., Central Experimental Farm, Ottawa, Ontario, KIA 0C6,
 CANADA
 Tel: (613) 759-1658
 Fax: (613) 759-1701
 Email: ouellet@agr.gc.ca.

FEATURES
 source
 1..511
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="Frontana"
 /db_xref="taxon:4565"
 /clone="Ta01_09G01"
 /tissue_type="heads"
 /dev_stage="anthesis"
 /clone_lib="Ta01_AAFc_ECOrc_Fusarium graminearum inoculate
 d_wheat_heads"
 /note="Vector: pGEM-T easy; Site_1: EcoRI; Site_2: EcoRI;
 Controlled chamber-grown wheat heads were spray inoculated
 at mid-anthesis with a Fusarium graminearum macroconidial
 suspension (50,000 spores/ml) and kept under intermittent
 misting for 24 hours, then collected and immediately
 frozen in liquid nitrogen."

ORIGIN
 Alignment Scores:
 Pred. No.: 24 Length: 511
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.64% Indels: 0
 DB: 4 Gaps: 0

US-10-069-353a-8 (1-275) x B1750759 (1-511)

QY 148 GluProAspArgAla1leuArgGlu1leu 157
 |||||
 Db 159 GAACCGAGACCGTCGATTCGGAGATTGG 130

RESULT 11
 LOCUS BQ620788/c 540 bp mRNA linear EST 28-JUN-2002
 DEFINITION Talr1111C01R Talr1 Triticum aestivum cDNA clone Talr1111C01R, mRNA
 sequence.
 ACCESSION BQ620788
 VERSION BQ620788.1 GI:21625867
 KEYWORDS EST
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.
 1 (bases 1 to 540)
 REFERENCE
 AUTHORS Cloutier, S., Dong, G. and Walsh, A.
 TITLE Wheat functional genomics - Thatcher Lr1 cDNA library
 JOURNAL Unpublished (2001)
 COMMENT Contact: Dr. Sylvie Cloutier
 Cereal Research Centre, Agriculture and Agri-food Canada
 195 D'Almeida Rd, Winnipeg, MB, Canada R3T 2M9
 Tel: (204) 983-2340
 Fax: (204) 983-4604
 Email: scoutier@agr.gc.ca
 was cloned directionally, not all sequences generated with reverse
 primer where from the 5' end (same with forward primer and 3' end).
 Average insert size is >2.2kb
 Plate: 111 row: C column: 01
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1..540

/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Thatcher Lr1"
/db_xref="taxon:4555"
/clone="TAB111C01R"
/tissue_type="Leaf tissue"
/dev_stage="14 Days old"
/lab_host="E. coli XL0R"
/clone_lib="TatLr1"
/note="Vector: lambda ZapII; mass excised in plasmid
vector pK-CMV (Stratagene). Site_1: EcoRI; Site_2: XhoI;
mRNA obtained from wheat NIL Thatcher Lr1 24 hours after
inoculation with leaf rust pathogen Puccinia triticina
race BBB carrying the avirulence gene Avr1."

ORIGIN

Alignment Scores:

Pred. No.:	25.2	Length:	540
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.64%	Indels:	0
DB:	5	Gaps:	0

US-10-069-353A-8 (1-275) x B0620788 (1-540)

Qy

148 GtProAspArgAlaIleArgGluIleLeu 157
|||||
230 GAGCCGACCGTCGATTCGGAGATTG 201

RESULT 12

BU552981 552 bp mRNA linear EST 14-NOV-2002
LOCUS BU552981 K. Sato unpublished cDNA library, strain H602 adult.
DEFINITION heading stage top three leaves Hordeum vulgare subsp. spontaneum
CDNA clone bahe2h02 3', mRNA sequence.

ACCESSION

BU552981 GI:24971432

VERSION

BU552981

KEYWORDS

EST.

SOURCE

Hordeum vulgare subsp. spontaneum

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Hordeum.
1 (bases 1 to 552)

REFERENCE

Sato, K., Saitoh, D. and Takeda, K.

AUTHORS

Barley EST sequencing project in NIG and Okayama Univ

TITLE

Unpublished (2002)

JOURNAL

Contact: Tadao Shin-i

COMMENT

Center for Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.

FEATURES

Location/Qualifiers

1..552

/organism="Hordeum vulgare subsp. spontaneum"

/mol_type="mRNA"

/strain="H602"

/sub_species="spontaneum"

/db_xref="taxon:77009"

/clone="bahe2h02"

/tissue_type="top three leaves"

/dev_stage="adult, heading stage"

/clone_lib="K. Sato unpublished cDNA library, strain H602
adult, heading stage top three leaves"

ORIGIN

Alignment Scores:
Pred. No.: 25.7 Length: 552
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.64% Indels: 0
DB: 4 Gaps: 0

US-10-069-353A-8 (1-275) x BU552981 (1-552)

Qy

148 GtProAspArgAlaIleArgGluIleLeu 157
|||||
201 GAGCCGACCGTCGATTCGGAGATTG 172

RESULT 13

CD876873 564 bp mRNA linear EST 11-JUL-2003
LOCUS CD876873 C AZ03.111C19R011127 AZ03 Triticum aestivum cDNA clone AZ03111C19,
DEFINITION mRNA sequence.

ACCESSION

CD876873 GI:32560689

VERSION

CD876873

KEYWORDS

EST.

SOURCE

Triticum aestivum (bread wheat)

ORGANISM

Triticum aestivum

REFERENCE

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Triticum.
1 (bases 1 to 564)

AUTHORS

Genoplatne.

TITLE

Genoplatne, a major partnership french program in plant genomics

JOURNAL

Unpublished (2003)

COMMENT

Contact: Genoplatne

Genoplatne
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplatne' (<http://www.genoplatne.com>
and <http://genoplatne-info.infobioigen.fr>).

FEATURES

Location/Qualifiers

1..564

/organism="Triticum aestivum"

/mol_type="mRNA"

/cultivar="rectal"

/db_xref="taxon:4565"

/clone="AZ03111C19"

/tissue_type="leaf"

/clone_lib="AZ03"

ORIGIN

Alignment Scores:
Pred. No.: 26.2 Length: 564
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.64% Indels: 0
DB: 6 Gaps: 0

US-10-069-353A-8 (1-275) x CD876873 (1-564)

Qy

148 GtProAspArgAlaIleArgGluIleLeu 157
|||||
229 GAGCCGACCGTCGATTCGGAGATTG 200

RESULT 14

BU459101 567 bp mRNA linear EST 23-MAY-2002
LOCUS BU459101 K. Sato unpublished cDNA library, cv. Akashinriki
DEFINITION vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone
baa34i03 3', mRNA sequence.

ACCESSION

BU459101 GI:21137635

VERSION

BU459101

KEYWORDS

EST.

SOURCE

Hordeum vulgare subsp. vulgare

ORGANISM

Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

REFERENCE Poideae; Triticaceae; Hordeum.
1 (bases 1 to 567)
AUTHORS Sato, K., Saisho, D. and Takeda, K.
TITLE Barley EST sequencing project in NIG and Okayama Univ
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..567
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Akashinriki"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="bak34103"
/issue_type="leaves"
/dev_stage="vegetative stage"
/clone_lib="K. Sato unpublished cDNA library, cv.
Akashinriki vegetative stage leaves"

ORIGIN

Alignment Scores:

Pred. No.:	26.4	Length:	567
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.64%	Indels:	0
DB:	4	Gaps:	0

US-10-069-353a-8 (1-275) x BJ459101 (1-567)

OY 154 ArgGluLeuValArgValLeuLysProGly 163
|||||
Db 372 CGAGAGATCTGAGAGTCTAAACACGCG 343
|||||

RESULT 15 572 bp mRNA linear EST 18-JAN-2002
AV945873/c LOCUS
AV945873 K. Sato unpublished cDNA library, strain H602 adult,
DEFINITION heading stage top three leaves Hordeum vulgare subsp. spontaneum
cDNA clone bah28k06 3', mRNA sequence.
AV945873
ACCESSION AV945873.1 GI:18241670
VERSION
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. spontaneum
ORGANISM Hordeum vulgare subsp. spontaneum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticaceae; Hordeum.
1 (bases 1 to 572)

REFERENCE Sato, K., Saisho, D. and Takeda, K.
TITLE Barley EST sequencing project in NIG and Okayama Univ
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..572
/organism="Hordeum vulgare subsp. spontaneum"
/mol_type="mRNA"
/strain="H602"
/sub_species="spontaneum"
/db_xref="taxon:77009"
/clone="bah28k06"

/issue_type="top three leaves"
/dev_stage="adult, heading stage"
/clone_lib="K. Sato unpublished cDNA library, strain H602
adult, heading stage top three leaves"

ORIGIN

Pred. No.:	26.6	Length:	572
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.64%	Indels:	0
DB:	1	Gaps:	0

US-10-069-353a-8 (1-275) x AV945873 (1-572)

OY 148 GIUPROSPARGA1a1leargluileu 157
|||||
Db 201 GAACCGATCGTGCATTCGGGAGATTTA 172
|||||

Search completed: August 30, 2005, 01:03:29
Job time : 3551 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 29, 2005, 22:25:40 ; Search time 227 Seconds
(without alignments)
1982.274 Million cell updates/sec

Title: US-10-069-353A-8

Perfect score: 275
Sequence: 1 VLPGAPTSQGVGMVLT.....YKVAHDMGALITARKPVG 275

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2400006

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-O/cgn2_1/USPTO.spool/US10069353/runat.26082005.172152.15080/app.query.fasta_1.455
-Db=Issued Patents.NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOFCG=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10069353 @CCN 1.1 69 @runat.26082005.172152.15080 -NCPU=6 -ICPU=3
-NO MAP -LARGEOUTERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -MARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents.NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	275	100.0	80161	3	US-09-036-987A-1
2	275	100.0	80161	3	US-09-370-700-1
3	275	100.0	80161	4	US-09-603-207-1
4	10	3.6	738	4	US-09-724-797-13
5	10	3.6	6085	3	US-09-029-603-4
6	9	3.3	400	4	US-08-956-171E-2979
7	9	3.3	400	4	US-08-781-986A-2979
8	9	3.3	11466	4	US-08-956-171E-444
9	9	3.3	11466	4	US-08-781-986A-444
10	8	2.9	310	4	US-09-270-767-28714
11	8	2.9	414	4	US-09-252-991A-334
12	8	2.9	456	4	US-09-248-796A-3885

C	13	8	2.9	505	4	US-09-621-976-18024	Sequence 18024, A	
	14	8	2.9	507	4	US-09-902-540-7484	Sequence 7484, Ap	
	15	8	2.9	522	4	US-09-902-540-3761	Sequence 3761, Ap	
	C	16	8	2.9	579	3	US-08-991-789A-147	Sequence 147, App
	C	17	8	2.9	579	3	US-09-062-451-147	Sequence 147, App
	C	18	8	2.9	579	3	US-09-598-326-147	Sequence 147, App
	C	19	8	2.9	579	4	US-09-289-198-147	Sequence 147, App
	C	20	8	2.9	579	4	US-09-429-755-147	Sequence 147, App
	C	21	8	2.9	579	4	US-09-699-295-147	Sequence 147, App
	C	22	8	2.9	597	4	US-09-252-991A-16384	Sequence 16384, A
	C	23	8	2.9	630	4	US-09-252-991A-16006	Sequence 16006, A
	C	24	8	2.9	651	4	US-09-252-991A-7735	Sequence 7735, Ap
	C	25	8	2.9	660	4	US-09-252-991A-16273	Sequence 16273, A
	C	26	8	2.9	759	4	US-09-583-110-637	Sequence 637, App
	27	8	2.9	768	4	US-09-902-540-5226	Sequence 5226, Ap	
	28	8	2.9	780	4	US-09-107-433-1646	Sequence 1646, Ap	
	C	29	8	2.9	789	4	US-09-252-991A-11034	Sequence 11034, A
	30	8	2.9	891	1	US-08-687-895-2	Sequence 2, Appl1	
	31	8	2.9	891	2	US-09-040-482-2	Sequence 2, Appl1	
	32	8	2.9	933	4	US-09-248-796A-6460	Sequence 6460, Ap	
	C	33	8	2.9	1101	4	US-09-489-039A-2638	Sequence 2638, Ap
	34	8	2.9	1110	4	US-09-252-991A-312	Sequence 312, Appl	
	35	8	2.9	1161	4	US-09-252-991A-7650	Sequence 7650, Ap	
	36	8	2.9	1161	4	US-09-270-767-12864	Sequence 12864, A	
	37	8	2.9	1212	4	US-09-489-039A-5791	Sequence 5791, Ap	
	38	8	2.9	1232	3	US-08-411-760-13	Sequence 13, Appl	
	39	8	2.9	1233	3	US-08-244-491A-1	Sequence 1, Appl1	
	40	8	2.9	1233	2	US-08-797-727A-1	Sequence 1, Appl1	
	41	8	2.9	1282	4	US-09-205-258-197	Sequence 197, Appl	
	42	8	2.9	1401	4	US-09-602-787A-67	Sequence 67, Appl	
	43	8	2.9	1424	4	US-09-270-767-9997	Sequence 9997, Ap	
	44	8	2.9	1692	4	US-09-902-540-4219	Sequence 4219, Ap	
	C	45	8	2.9	1719	4	US-09-252-991A-353	Sequence 353, Appl

ALIGNMENTS

RESULT 1
US-09-036-987A-1
; Sequence 1, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Crawford, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Madduri, Donald J.
; APPLICANT: Treadway, Patli J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; TITLE OF INVENTION: Production
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow Agrosciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608


```
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (317)337-4816
/ TELEFAX: (317)337-4847
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 80161 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
US-09-036-987A-1

Alignment Scores:
Pred. No.: 5,22e-273 Length: 80161
Score: 275.00 Matches: 275
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 3

US-10-069-353A-8 (1-275) x US-09-036-987A-1 (1-80161)

QY 1 ValLeuProGlyValAlaProThrSerGlnGlnValAlaGlyMetTyrAspLeuValThr 20
Db 20168 GTGTGGCCAGGTGGGCGCACCAATCGCAGGTTGGGCAATGATGACCTGTGTCAG 20227
QY 21 ProLeuLeuAsnSerValAlaGlyGlyProCysAlaIleHisGlyTyrTrpGluAsn 40
Db 20228 CCGTTCCTGAATCTCGGTCCGGGCGGGCCCTTCGCCATTCACACCGCTACTGGAGAAC 20287
QY 41 AspGlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGlyArg 60
Db 20288 GACGGGCGGGCTTCTCGGCGAGCGCGCCGACCGGCTCACCGACTTGTCCGCAACGG 20347
QY 61 ThrValLeuAspGlyValAlaArgLeuLeuAspValAlaGlyCysGlyThrGlyGlnProAla 80
Db 20348 ACCGAGCTCGATGGCGGCTGTGACTGTGATGTGGGGTGGGTACCGGACCAACCAACG 20407
QY 81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGln 100
Db 20408 CTGGCGCTCGCGCGGACCAACCGCATCCGATCACCGGCTACCGCTGACCGAGTGCA 20467
QY 101 ValAlaIleAlaAlaAspCysAlaArgGlyLeuArgGlyLeuSerHisArgValAspPheSer 120
Db 20468 GTGGCGATCGCGCGATGTCGACGCGACCGGACCTPAAGCCACCGGGTGGACTTTCG 20527
QY 121 CysValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
Db 20528 TCGCTCGATGCGATGTCTCTGCGGACCGGACCAATGCTTTCGACGCGCTGGGCGCATG 20587
QY 141 GlnSerLeuLeuGlnMetSerGluProAspArgAlaIleArgGlyIleLeuArgValLeu 160
Db 20588 CAGTGGCTGTGGAGATGTCCCAACCGACCTGCGATCCGGGAATCTTCTGATGACTC 20647
QY 161 LysProGlyGlyIleLeuGlyValThrGlyValValLeuArgGlyAlaGlyIleGlyMet 180
Db 20648 AAACCCGGGCGATCTCGGCGCTCACCGAGTGTCTCAAAACGAGAGCGGGCGCGGATG 20707
QY 181 ProValSerGlyAspArgTrpProThrArgIleLeuArgIleCysLeuAlaGlnGlnLeu 200
Db 20708 CCGGTCGTCGGGGACAGGTGGCGACCGGCTTCGATCTGCTCGAGAACCTTTCG 20767
QY 201 GlnSerLeuArgAlaAlaGlyPheGlyIleLeuAspTrpGluAspValSerSerArgThr 220
Db 20768 GAATGCTGCTCGCGACGGGGTTCCAGATCTTCGATTTGGAGAGACGTGTCTCGAGGACC 20827
QY 221 ArgTyrPheMetProGlnPheAlaGlnGlnLeuAlaAlaHisGlnHisGlyIleAlaAsp 240
Db 20828 CGGACTTCATGCCCGACATTCGCGCAAGAGCTCGCTGGCGACACACACGCGGATCCGGAC 20887
QY 241 ArgTyrArgIleProAlaValAlaGlyTyrAlaAlaAlaValCysAspTyrGlyLeuArgTyrAla 260
Db 20888 AGGTACGGGCGCGCTGTGCTGGCGGCGCGGCTGTGCGATTATGAGAAATATGCC 20947
```

```
QY 261 HisAspMetGlyTyrAlaIleLeuThrAlaArgIleYspProValGly 275
Db 20948 CACGACATGGGCTATGCGATTCTGACGGCGGGAAGCCGCTCGGC 20992

RESULT 2
US-09-370-700-1
/ Sequence 1, Application US/09370700
/ Patent No. 6274350
/ GENERAL INFORMATION:
/ APPLICANT: Balcz, Richard H
/ APPLICANT: Broughton, Mary C
/ APPLICANT: Crawford, Kathryn P
/ APPLICANT: Madduri, Krishnamurthy
/ APPLICANT: Treadway, Patci J
/ APPLICANT: Turner, Jan R
/ APPLICANT: Waldron, Clive
/ TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
/ FILE REFERENCE: 50489 DIV1
/ CURRENT FILING DATE: 1999-08-09
/ EARLIER FILING DATE: 1998-03-09
/ NUMBER OF SEQ ID NOS: 39
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 1
/ LENGTH: 80161
/ TYPE: DNA
/ ORGANISM: Saccharopolyspora spinosa
US-09-370-700-1

Alignment Scores:
Pred. No.: 5,22e-273 Length: 80161
Score: 275.00 Matches: 275
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 3

US-10-069-353A-8 (1-275) x US-09-370-700-1 (1-80161)

QY 1 ValLeuProGlyValAlaProThrSerGlnGlnValAlaGlyMetTyrAspLeuValThr 20
Db 20168 GTGTGGCCAGGTGGGCGCACCAATCGCAGGTTGGGCAATGATGACCTGTGTCAG 20227
QY 21 ProLeuLeuAsnSerValAlaGlyGlyProCysAlaIleHisGlyTyrTrpGluAsn 40
Db 20228 CCGTTCCTGAATCTCGGTCCGGGCGGGCCCTTCGCCATTCACACCGCTACTGGAGAAC 20287
QY 41 AspGlyArgAlaSerTrpGlnGlnAlaAlaAspArgLeuThrAspLeuValAlaGlyArg 60
Db 20288 GACGGGCGGGCTTCTCGGCGAGCGCGCCGACCGGCTCACCGACTTGTGCGCAACGG 20347
QY 61 ThrValLeuAspGlyValAlaArgLeuLeuAspValAlaGlyCysGlyThrGlyGlnProAla 80
Db 20348 ACCGAGCTCGATGGCGGCTGTGACTGTGATGTGGGGTGGGTACCGGACCAACGCG 20407
QY 81 LeuArgValAlaArgAspAsnAlaIleGlnIleThrGlyIleThrValSerGlnValGln 100
Db 20408 CTGGCGCTCGCGCGGACCAACCGCATCCGATCACCGGCTACCGCTGACCGAGTGCA 20467
QY 101 ValAlaIleAlaAlaAspCysAlaArgGlyLeuArgGlyLeuSerHisArgValAspPheSer 120
Db 20468 GTGGCGATCGCGCGATGTCGACGCGGACCGGACCTPAAGCCACCGGGTGGACTTTCG 20527
QY 121 CysValAspAlaMetSerLeuProTyrProAspAsnAlaPheAspAlaAlaTrpAlaMet 140
Db 20528 TCGCTCGATGCGATGTCTCTGCGGACCGGACCAATGCTTTCGACGCGCTGGGCGCATG 20587
QY 141 GlnSerLeuLeuGlnMetSerGluProAspArgAlaIleArgGlyIleLeuArgValLeu 160
Db 20588 CAGTGGCTGTGGAGATGTCCCAACCGGACCGGCTTCGATCCGGGAATCTTCTGAGTACTC 20647
```


COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2979:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-2979

Alignment Scores:
Pred. No.: 3.89 Length: 400
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 4 Gaps: 0

US-10-069-353a-8 (1-275) x US-08-781-986A-2979 (1-400)

QY 158 ArgValIleuLysProGlyGlyIleu 166
Db 151 CGAGCTTAAAGCTGGGGGTATCTTG 125

RESULT 8
US-08-956-171E-444
Sequence 444, Application US/08956171E
Patent No. 6593114
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
Gill H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996

APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 444:
SEQUENCE CHARACTERISTICS:
LENGTH: 11466 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-956-171E-444

Alignment Scores:
Pred. No.: 94.9 Length: 11466
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 4 Gaps: 0

US-10-069-353a-8 (1-275) x US-08-956-171E-444 (1-11466)

QY 158 ArgValIleuLysProGlyGlyIleu 166
Db 10859 CGAGCTTAAAGCTGGGGGTATCTTG 10885

RESULT 9
US-08-781-986A-444
Sequence 444, Application US/08781986A
Patent No. 6737248
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 444:
SEQUENCE CHARACTERISTICS:
LENGTH: 11466 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-444

Alignment Scores:

Pred. No.:	94.9	Length:	11466
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.27%	Indels:	0
DB:	4	Gaps:	0

US-10-069-353A-8 (1-275) x US-08-781-986A-444 (1-11466)

QY 158 ArgValLeuValProGlyGlyIleu 166
DB 10859 CGAGTCTAAGCCTGGGGGATCTTG 10885

RESULT 10

US-09-270-767-28714
; Sequence 28714, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 28714
; LENGTH: 310
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-28714

Alignment Scores:

Pred. No.:	33	Length:	310
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.91%	Indels:	0
DB:	4	Gaps:	0

US-10-069-353A-8 (1-275) x US-09-270-767-28714 (1-310)

QY 70 LeuAspValGlyCysGlyThrGly 77
DB 48 TTGGATGTGGATGCGGCA CGGGA 71

RESULT 11

US-09-252-991A-334/C
; Sequence 334, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 334
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-334

Alignment Scores:

Pred. No.:	43.5	Length:	414
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.91% Indels: 0
DB: 4 Gaps: 0

US-10-069-353A-8 (1-275) x US-09-252-991A-334 (1-414)

QY 157 LeuArgValLeuValProGlyGly 164
DB 44 CTGAGGTCCTCAAGCCGCGGC 21

RESULT 12

US-09-248-796A-3885
; Sequence 3885, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 3885
; LENGTH: 456
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-3885

Alignment Scores:

Pred. No.:	47.7	Length:	456
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.91%	Indels:	0
DB:	4	Gaps:	0

US-10-069-353A-8 (1-275) x US-09-248-796A-3885 (1-456)

QY 70 LeuAspValGlyCysGlyThrGly 77
DB 199 CTTGATGTTGATGTGTA CTGT 222

RESULT 13

US-09-621-976-18024/C
; Sequence 18024, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 18024
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-18024

Alignment Scores:

Pred. No.:	52.6	Length:	505
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.91%	Indels:	0
DB:	4	Gaps:	0

US-10-069-353A-8 (1-275) x US-09-621-976-18024 (1-505)

QY 244 ProAlaValAlaGlyTrpAlaAla 251
DB 296 CCGGACGTGGCGGCTGGGACAGCA 273

RESULT 14
US-09-902-540-7484

; Sequence 7484, Application US/09902540

; Patent No. 6833447

; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/09/902,540

; PRIOR FILING DATE: 2001-07-10

; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 16825

; SEQ ID NO 7484

; LENGTH: 507

; TYPE: DNA

; ORGANISM: Myxococcus xanthus

US-09-902-540-7484

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-069-353A-8 (1-275) x US-09-902-540-7484 (1-507)

QY 70 LeuAspValAlaGlyCysGlyThrGly 77

DB 28 TTGGACGTGGCGGCTGGGACAGCGG 51

RESULT 15
US-09-902-540-3761

; Sequence 3761, Application US/09902540

; Patent No. 6833447

; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/09/902,540

; PRIOR FILING DATE: 2001-07-10

; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 16825

; SEQ ID NO 3761

; LENGTH: 522

; TYPE: DNA

; ORGANISM: Myxococcus xanthus

US-09-902-540-3761

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-069-353A-8 (1-275) x US-09-902-540-3761 (1-522)

QY 70 LeuAspValAlaGlyCysGlyThrGly 77

DB 58 CTGGACGTGGCGGCTGGGACAGCGG 81

Search completed: August 30, 2005, 01:08:17
Job time : 288 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 29, 2005, 22:45:45 ; Search time 699 Seconds

(without alignments)
2574.177 Million cell updates/sec

Title: US-10-069-353a-8

Perfect score: 275

Sequence: 1 VLPGGAPTSQQVGMIDLVY.....YKRYADMGYAILTARKPVG 275

Scoring table:

OLIGO	Xgapop 60.0 , Xgapext 60.0
	Ygapop 60.0 , Ygapext 60.0
	Fgapop 6.0 , Fgapext 7.0
	Delop 6.0 , Delext 7.0

Searched: 7331713 segs, 3271544945 residues

Word size: 1

Total number of hits satisfying chosen parameters: 14651541

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

```
-MODE=frame+ p2n.model -DEV=x1h
-O/cgn2_1/USPTO.spool/US10069353.rnuc 26082005 172153 15107/app.query.fasta_1.455
-DB=Published Applications NA -QFMT=fastap -SUPFLX=rmdb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cgi -LIST=45 -DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEXPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10069353 @cgn 1 1 480 @rnuc 26082005 172153 15107
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -MAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7
```

Database : Published Applications NA:*

```
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
3: /cgn2_6/prodata/1/pubpna/US05_PUBCOMB.seq:*
4: /cgn2_6/prodata/1/pubpna/US04_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US03_PUBCOMB.seq:*
6: /cgn2_6/prodata/1/pubpna/US02_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US01_PUBCOMB.seq:*
8: /cgn2_6/prodata/1/pubpna/US00_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
11: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
12: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
13: /cgn2_6/prodata/1/pubpna/US05_PUBCOMB.seq:*
14: /cgn2_6/prodata/1/pubpna/US04_PUBCOMB.seq:*
15: /cgn2_6/prodata/1/pubpna/US03_PUBCOMB.seq:*
16: /cgn2_6/prodata/1/pubpna/US02_PUBCOMB.seq:*
17: /cgn2_6/prodata/1/pubpna/US01_PUBCOMB.seq:*
18: /cgn2_6/prodata/1/pubpna/US00_PUBCOMB.seq:*
19: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
20: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
21: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
22: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
23: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
24: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
25: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*
26: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	275	100.0	80161	17	US-10-329-148A-1	Sequence 1, Appl1
2	9	3.3	400	8	US-08-781-986A-2979	Sequence 2979, Ap
3	9	3.3	400	18	US-10-329-624-2979	Sequence 2979, Ap
4	9	3.3	11466	8	US-08-781-986A-444	Sequence 444, App
5	9	3.3	11466	18	US-10-329-624-444	Sequence 444, App
6	9	3.3	47745	13	US-10-087-132-469	Sequence 469, App
7	9	3.3	64492	18	US-10-378-083-1	Sequence 1, Appl1
8	9	2.9	262	20	US-10-425-115-21054	Sequence 21054, A
9	8	2.9	270	9	US-09-954-456-1002	Sequence 1002, Ap
10	8	2.9	270	10	US-09-960-706-835	Sequence 835, App
11	8	2.9	270	10	US-09-873-319-539	Sequence 539, App
12	8	2.9	270	17	US-10-170-385-168	Sequence 168, App
13	8	2.9	270	21	US-10-843-641A-4029	Sequence 4029, App
14	8	2.9	318	20	US-10-425-115-55334	Sequence 55334, A
15	8	2.9	333	15	US-10-156-761-81	Sequence 81, Appl
16	8	2.9	340	20	US-10-425-115-154357	Sequence 154357, A
17	8	2.9	344	20	US-10-425-115-138703	Sequence 138703, A
18	8	2.9	349	20	US-10-357-930-7576	Sequence 9363, Ap
19	8	2.9	381	20	US-10-425-115-9363	Sequence 9363, Ap
20	8	2.9	405	20	US-10-357-930-37536	Sequence 37536, A
21	8	2.9	441	21	US-10-472-928-2613	Sequence 2613, Ap
22	8	2.9	459	19	US-10-474-776-518	Sequence 518, App
23	8	2.9	483	10	US-09-918-995-617	Sequence 617, App
24	8	2.9	484	10	US-09-918-995-29551	Sequence 29551, A
25	8	2.9	513	18	US-10-424-599-130420	Sequence 130420, A
26	8	2.9	514	9	US-09-920-300A-1537	Sequence 1537, Ap
27	8	2.9	514	13	US-10-033-528-1637	Sequence 1637, Ap
28	8	2.9	514	16	US-10-099-926-1637	Sequence 1637, Ap
29	8	2.9	514	22	US-10-961-527-1637	Sequence 1637, Ap
30	8	2.9	528	17	US-10-259-194A-625	Sequence 625, App
31	8	2.9	561	13	US-10-027-632-32254	Sequence 32254, A
32	8	2.9	561	17	US-10-027-632-32254	Sequence 32254, A
33	8	2.9	564	19	US-10-767-701-29067	Sequence 29067, A
34	8	2.9	576	18	US-10-424-599-138183	Sequence 138183, A
35	8	2.9	579	9	US-09-810-936-147	Sequence 147, App
36	8	2.9	579	9	US-09-429-755-147	Sequence 147, App
37	8	2.9	579	9	US-09-924-400-147	Sequence 147, App
38	8	2.9	579	15	US-10-212-679-147	Sequence 147, App
39	8	2.9	579	18	US-10-079-137B-147	Sequence 147, App
40	8	2.9	599	22	US-10-972-079-31973	Sequence 31973, A
41	8	2.9	600	22	US-10-972-079-3372	Sequence 3272, Ap
42	8	2.9	600	22	US-10-972-079-3373	Sequence 3273, Ap
43	8	2.9	600	22	US-10-972-079-3274	Sequence 3274, Ap
44	8	2.9	600	22	US-10-972-079-3275	Sequence 3275, Ap
45	8	2.9	600	22	US-10-972-079-31974	Sequence 31974, A

ALIGNMENTS

```
RESULT 1
US-10-329-148A-1
: Sequence 1, Application US/10329148A
: Publication No. US20040023343A1
:
GENERAL INFORMATION:
: APPLICANT: Baltz, Richard H
: APPLICANT: Broughton, Mary C
: APPLICANT: Crawford, Kathryn P
: APPLICANT: Madduri, Krishnamurthy
: APPLICANT: Treadway, Patil J
: APPLICANT: Turner, Jan R
: APPLICANT: Waldron, Clive
:
: FILE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
: FILE REFERENCE: 50489 DIV1
: CURRENT APPLICATION NUMBER: US/10/329,148A
: CURRENT FILING DATE: 2002-12-23
```


NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/329,624
FILING DATE: 27-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/956,171
FILING DATE: October 20, 1997
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2979:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2979:
US-10-329-624-2979
Alignment Scores:
Pred. No.: 19.5 Length: 400
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 18 Gaps: 0
US-10-069-353a-8 (1-275) x US-10-329-624-2979 (1-400)
QY 158 ArgValIeuIySProGIyGIyIleIeu 166
Db 151 CGAGTCTTAAAGCTGGGGGTATCTTG 125
RESULT 4
US-08-781-986A-444
Sequence 444, Application US/08781986A
Publication No. US2003005436A1
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 444:
SEQUENCE CHARACTERISTICS:
LENGTH: 11466 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-444
Alignment Scores:
Pred. No.: 339 Length: 11466
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 8 Gaps: 0
US-10-069-353a-8 (1-275) x US-08-781-986A-444 (1-11466)
QY 158 ArgValIeuIySProGIyGIyIleIeu 166
Db 10859 CGAGTCTTAAAGCTGGGGGTATCTTG 10885
RESULT 5
US-10-329-624-444
Sequence 444, Application US/10329624
Publication No. US20040043037A1
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
Gil H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/329,624
FILING DATE: 27-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/956,171
FILING DATE: October 20, 1997
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman

```

;
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248PID1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 444:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 11466 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 444:
US-10-329-624-444

Alignment Scores:
Pred. No.: 339 Length: 11466
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 18 Gaps: 0

US-10-069-353A-8 (1-275) x US-10-329-624-444 (1-11466)

QY 158 ArgValAlaGluArgThrValLeuasp 166
DB 10859 CGAGCTTTAAAGCTGGGGGTATCTTG 10885

RESULT 6
US-10-087-192-469/c
; Sequence 469, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 469
; LENGTH: 47745
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(47745)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-469

Alignment Scores:
Pred. No.: 1.14e+03 Length: 47745
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 13 Gaps: 0

US-10-069-353A-8 (1-275) x US-10-087-192-469 (1-47745)

QY 56 LeuValAlaGluArgThrValLeuasp 64
DB 35691 TTAGTGGCAGAGAGACTGTACTTAC 35665

RESULT 7
US-10-378-083-1/c
; Sequence 1, Application US/10378083
```

```

; Publication No. US20040053274A1
; GENERAL INFORMATION:
; APPLICANT: President of Tokyo Institute of Technology
; TITLE OF INVENTION: A gene cluster of vicenistatin biosynthase, a vicenistamine
; glycosyltransferase polypeptide, and a gene encoding the
; TITLE OF INVENTION: polypeptide
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/378,083
; CURRENT FILING DATE: 2003-03-04
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 64492
; TYPE: DNA
; ORGANISM: Streptomyces halstedii sp.HC-34
US-10-378-083-1

Alignment Scores:
Pred. No.: 1.47e+03 Length: 64492
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 18 Gaps: 0

US-10-069-353A-8 (1-275) x US-10-378-083-1 (1-64492)

QY 69 LeuLeuAspValGlyCysGlyThrGly 77
DB 33852 CTGCTCGATGTGGGCTGTGGCACCGGC 33826

RESULT 8
US-10-425-115-21054/c
; Sequence 21054, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 21054
; LENGTH: 262
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(262)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: WRT4577_119199C.1
US-10-425-115-21054

Alignment Scores:
Pred. No.: 145 Length: 262
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.91% Indels: 0
DB: 20 Gaps: 0

US-10-069-353A-8 (1-275) x US-10-425-115-21054 (1-262)

QY 63 LeuAspGlyValArgLeuLeu 70
DB 47 CTGACGGGGGGGTTCGCTCTC 24

RESULT 9
```

```
US-09-954-456-1002/c
; Sequence 1002, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; FILE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1002
; LENGTH: 270
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: n=a,t,g or c
US-09-954-456-1002

Alignment Scores:
Pred. No.: 149 Length: 270
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.91% Indels: 0
DB: 9 Gaps: 0

US-10-069-353a-8 (1-275) x US-09-954-456-1002 (1-270)

QY 17 AspleuValThProleuLeuAn 24
Db 92 GATCTGTCACCTCTTGTCTCAAC 69

RESULT 10
US-09-960-706-835/c
; Sequence 835, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplas
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 835

; LENGTH: 270
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(270)
; OTHER INFORMATION: n = a or c or g or t
US-09-960-706-835

Alignment Scores:
Pred. No.: 149 Length: 270
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.91% Indels: 0
DB: 10 Gaps: 0

US-10-069-353a-8 (1-275) x US-09-873-319-539 (1-270)

QY 17 AspleuValThProleuLeuAn 24
Db 92 GATCTGTCACCTCTTGTCTCAAC 69

RESULT 12
US-10-170-385-168/c
; Sequence 168, Application US/10170385
; Publication No. US2003020372A1
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Ward, Neil Raymond
/ APPLICANT: Mundy, Christopher Robert
/ APPLICANT: Kan, On
/ APPLICANT: Harris, Robert Alan
/ APPLICANT: White, Jonathan
/ APPLICANT: Binley, Katie Mary
/ APPLICANT: Rayner, William Nigel
/ APPLICANT: Naylor, Stuart
/ APPLICANT: Kingsman, Susan Mary
/ APPLICANT: Krige, David
/ TITLE OF INVENTION: ANALYSIS METHOD
/ FILE REFERENCE: 532682000100
/ CURRENT APPLICATION NUMBER: US/10/170,385
/ CURRENT FILING DATE: 2002-06-12
/ PRIOR APPLICATION NUMBER: PCT/GB02/01662
/ PRIOR FILING DATE: 2002-04-08
/ PRIOR APPLICATION NUMBER: PCT/GB01/05458
/ PRIOR FILING DATE: 2001-12-10
/ NUMBER OF SEQ ID NOS: 549
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 168
/ LENGTH: 270
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: 144..233, 248
/ OTHER INFORMATION: n = A,T,C or G
US-10-170-385-168
```

```
Alignment Scores:
Pred. No.: 149 Length: 270
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.91% Indels: 0
DB: 17 Gaps: 0
```

US-10-069-353A-8 (1-275) x US-10-170-385-168 (1-270)

```
OY 17 AspleuValThProLeuLeuAsn 24
|||
Db 92 GATCTCGTCACTCCTTGTCTCAAC 69
```

```
RESULT 13
US-10-843-641A-4029/C
/ Sequence 4029, Application US/10843641A
/ Publication No. US20050064454A1
/ GENERAL INFORMATION:
/ APPLICANT: Avalon Pharmaceuticals, Inc.
/ TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
/ FILE REFERENCE: 689290-189
/ CURRENT APPLICATION NUMBER: US/10/843,641A
/ CURRENT FILING DATE: 2004-05-12
/ PRIOR APPLICATION NUMBER: US/09/873,367
/ PRIOR FILING DATE: 2001-06-05
/ PRIOR APPLICATION NUMBER: US/09/954,531
/ PRIOR FILING DATE: 2001-09-18
/ PRIOR APPLICATION NUMBER: US/09/954,456
/ PRIOR FILING DATE: 2001-09-25
/ PRIOR APPLICATION NUMBER: US/09/962,436
/ PRIOR FILING DATE: 2001-09-25
/ PRIOR APPLICATION NUMBER: US/09/962,832
/ PRIOR FILING DATE: 2001-09-25
/ PRIOR APPLICATION NUMBER: US/09/964,824
/ PRIOR FILING DATE: 2001-09-27
/ PRIOR APPLICATION NUMBER: US/09/967,768
/ PRIOR FILING DATE: 2001-09-28
/ PRIOR APPLICATION NUMBER: US/09/968,007
/ PRIOR FILING DATE: 2001-10-02
/ PRIOR APPLICATION NUMBER: US/09/969,347
```

```
/ PRIOR FILING DATE: 2001-10-02
/ PRIOR APPLICATION NUMBER: US/09/969,708
/ PRIOR FILING DATE: 2001-10-03
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 8447
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 4029
/ LENGTH: 270
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (1)..(270)
/ OTHER INFORMATION: n=a,t,g or c
US-10-843-641A-4029
```

```
Alignment Scores:
Pred. No.: 149 Length: 270
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.91% Indels: 0
DB: 21 Gaps: 0
```

US-10-069-353A-8 (1-275) x US-10-843-641A-4029 (1-270)

```
OY 17 AspleuValThProLeuLeuAsn 24
|||
Db 92 GATCTCGTCACTCCTTGTCTCAAC 69
```

```
RESULT 14
US-10-425-115-55334
/ Sequence 55334, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 55334
/ LENGTH: 318
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: WRT4577_150460C.1
US-10-425-115-55334
```

```
Alignment Scores:
Pred. No.: 171 Length: 318
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.91% Indels: 0
DB: 20 Gaps: 0
```

US-10-069-353A-8 (1-275) x US-10-425-115-55334 (1-318)

```
OY 162 ProgiygiyleLeuGlyValThr 169
|||
Db 208 CCCGCGGTATCTGGGTGTAAAC 231
```

```
RESULT 15
US-10-156-761-81
/ Sequence 81, Application US/10156761
/ Publication No. US20030119018A1
/ GENERAL INFORMATION:
/ APPLICANT: OMURA, SATOSHI
```

APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 81
LENGTH: 333
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(333)
US-10-156-761-81

Alignment Scores:
Pred. No.: 178 Length: 333
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.91% Indels: 0
DB: 15 Gaps: 0

US-10-069-353a-8 (1-275) x US-10-156-761-81 (1-333)

Qy 70 LeuAspValGlyCysGlyThrGly 77
Db 295 CTGGACCTGGGCTGCCGACGGC 318

Search completed: August 30, 2005, 01:19:56
Job time : 748 secs

This Page Blank (uspto)